PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : C07K 13/00

A1

(11) International Publication Number: WO 92/10518

(43) International Publication Date: 25 June 1992 (25.06.92)

(21) International Application Number: PCT/US91/09055

(22) International Filing Date: 27 November 1991 (27.11.91)

(30) Priority data:
624,135 7 December 1990 (07.12.90) US

(71) Applicant: YALE UNIVERSITY [US/US]; 246 Church Street, New Haven, CT 06510 (US).

(71)(72) Applicants and Inventors: ROTHBERG, Jonathan, Marc [US/US]; 642 Rosemount Lane, Westville, CT 06515 (US). ARTAVANIS-TSAKONAS, Spyridon [US/US]; 192 Ridgewood Avenue, Hamden, CT 06517 (US).

(74) Agent: BARTH, Richard, S.; Frishauf, Holtz, Goodman & Woodward, 600 Third Avenue, New York, NY 10016 (US).

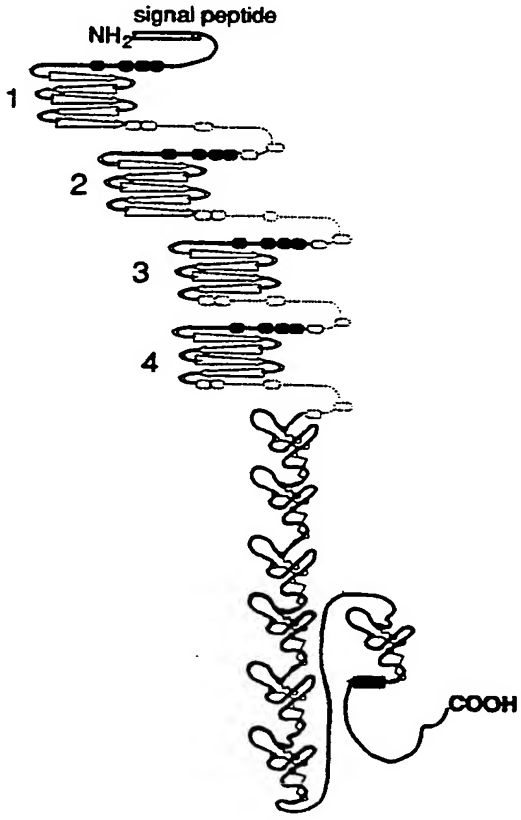
(81) Designated States: AT (European patent), BE (European patent), CA, CH (European patent), DE (European patent), DK (European patent), ES (European patent), FR (European patent), GB (European patent), GR (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent).

Published
With international search report.

(54) Title: PURIFIED SLIT PROTEIN AND SEQUENCE ELEMENTS THEREOF

(57) Abstract

An isolated and substantially pure form of the SLIT protein and sequence elements thereof, antibodies thereto and diagnostics and therapeutics utilizing such proteins and antibodies. A method for treating neurodegenerative disease, traumatic injury to a neural tissue or affecting the angiogenic process in a patient comprising administering to the patient an effective amount of the SLIT protein.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MG	Madagascar
AU	Australia	FI	Finland	ML	Mali
BB	Barbados	FR	France	MN	Mongolia
BE	Belgium	GA	Gabon	MR	Mauritania
BF	Burkina Faso	GB	United Kingdom	MW	Malawi
BG	Bulgaria	GN	Guinea	NL	Netherlands
BJ	Benin	GR	Greece	NO	Norway
BR	Brazil	HU	Hungary	PL	Poland
CA	Canada	IT	Italy	RO	Romania
CF	Central African Republic	JP	Japan	SD	Sudan
CG	Congo	KP	Democratic People's Republic of Korea	SE	Sweden
CH	Switzerland			SN	Senegal
CI	Côte d'Ivoire	KR	Republic of Korea	SU+	Soviet Union
CM	Cameroon	LI	Liechtenstein	TD	Chad
CS	Czechoslovakia	LK	Sri Lanka	TG	Togo
DE*	Germany	LU	Luxembourg	US	United States of America
DK	Denmark	MC	Monaco		

+ **Any designation of "SU" has effect in the Russian Federation. It is not yet known whether any such designation has effect in other States of the former Soviet Union.**

PURIFIED SLIT PROTEIN AND SEQUENCE ELEMENTS THEREOF

GOVERNMENT RIGHTS

This invention was made with United States government support under Grant NS 26084 from the National Institute of Health. The United States government thus has certain rights in this invention.

BACKGROUND OF THE INVENTION

Field of the Invention

The present invention concerns an isolated and substantially pure form of the SLIT protein and sequence elements thereof, antibodies thereto and diagnostics and therapeutics utilizing such proteins and antibodies.

Background Information

Proteins containing epidermal growth factor (EGF)-like sequences have been shown to play an important role in many aspects of eukaryotic cell control, acting as signals for proliferation, growth inhibition, and differentiation. A common feature of these proteins is their involvement in extracellular events and ligand-receptor interactions. In characterizing genomic DNA identified by cross-hybridization to the sequence coding for the tandem EGF-repeats of Notch, a gene involved in Drosophila neurogenesis, the isolation and partial characterization of sequences from an unlinked locus that coded for EGF-repeats have previously been reported. This sequence was shown to correspond to the SLIT locus and it was established that null mutations result in disruptions

of the embryonic CNS. (Rothberg, J.M., Hartley, D.A., Walther, Z., Artavanis-Tsakonas, S., (1988). slit: An EGF-Homologous Locus of D. Melanogaster Involved in the Development of the Embryonic Central Nervous System. Cell 55, 1047-1059).

The involvement of SLIT in cell interaction events is suggested by the presence of EGF-like repeats in the deduced protein sequence. Furthermore, both in situ hybridization, as well as antibody staining of embryos demonstrated that the highest level of slit expression is restricted to a special group of six midline glial cells that interact with and later enwrap developing commissural axons. Together, these findings are of particular interest, given the mutant phenotype and the evidence that, in both vertebrates and invertebrates, glial cells participate in neutral outgrowth through cell-cell contact and the secretion of diffusible factors (Bastiani, M. J., and Goodman, C.S. (1986). Guidance of neuronal growth cones in grasshopper embryo. III. Recognition of specific glial pathways. J. Neurosci. 6,3542-3551, reviewed in Vernadakis, A.(1988). Neuron-Glia Interrelations. Intern. Rev. Neurobiol., 30, 149-224).

The appearance of a glial scaffold in Drosophila before axonal outgrowth as well as the extension of pioneer growth cones along the surfaces of these glial cells, suggests that these glia play an instructive role in the determination of the major axon pathways in the central nervous system (CNS) (Jacob, J.R., and Goodman, C.S. (1989). Embryonic development of axon pathways in the drosophila CNS. I.A glial scaffold appears before the first growth cones. J. Neurosci. 9, 2402-2411; Jacobs, J.R., and Goodman, C.S. (1989). Embryonic development of axon pathways in the drosophila CNS. II. Behavior of pioneer growth cones. J. Neurosci. 9, 2402-2411).

It has long been thought that the extracellular environment influences the regulation of gene expression and the morphogenesis of cells during embryonic development (McDonald, J.A. (1989). Matrix regulation of cell shape and gene expression. Current Opinion in Cell Biology 1,995-999). In the nervous system, the morphogenetic events accompanying the formation of early structures have been shown to be dependent on the properties of the molecules that form their extracelular environment (see Jessell, (1988) Neuron. 1, 3-13). In vitro and in vivo studies suggest that growth cone guidance and axonal pathway selection are influenced by adhesive interactions between axons and extracellular matrix molecules (Sanes, J.R. (1989). Extracellular matrix molecules that influence neutral development. Ann. Rev. Neurosci. 12, 491-516).

Furthermore, specific constituents of the extracellular environment have been shown to affect neurite outgrowth in vitro and have been detected in vivo in the developing central and peripheral nervous systems (see Rutishauser, (1989), Neural cell-to-cell adhesion and recognition Current Opinion in Cell Biology, 1, 898-904).

Amino Acid Codes Amino Acid	Single Letter Code	Three Letter
Code		
alanine	A	Ala
cysteine	C	Cys
aspartic acid	D	Asp
glutamic acid	E	Glu
phenylalanine	F	Phe
glycine	G	Gly
histidine	H	His
isoleucine	I	Ile
lysine	K	Lys
leucine	L	Leu
methionine	M	Met
asparagine	N	Asn
proline	P	Pro
glutamine	Q	Gln
arginine	R	Arg
serine	S	Ser
threonine	T	Thr
valine	V	Val
tryptophan	W	Trp
tyrosine	Y	Tyr
any amino acid	X	

SUMMARY OF THE INVENTION

The present invention relates to recombinant proteins produced using all or part of the SLIT DNA sequences and exhibiting SLIT-like properties. The invention is also directed to the corresponding recombinant constructs and probes, including, genomic, cDNA, and synthetic DNA and protein sequences, as well as antibodies generated against specific domains of the SLIT protein. The invention also concerns prokaryotic and eukaryotic expression of all or parts of the SLIT-like genes from metazoan organisms, including, but not limited to its Flank-LRR-Flank and epidermal growth factor like sequences.

More specifically, the present invention concerns an isolated and substantially pure form of the SLIT protein comprising SEQ.ID. NO. 2, obtained by recombinant means from SEQ. ID. NO. 1 or from a natural source. The invention also relates to an isolated DNA segment encoding the entire SLIT protein, a recombinant expression vector comprising such DNA segment and a recombinant host microorganism containing a DNA expression vector comprising a DNA sequence consisting essentially of a DNA sequence encoding the entire SLIT protein.

The present invention also concerns a concensus amino flank-LRR-carboxy-flank sequence element of the SLIT protein (SEQ. I.D. NO. 8) comprising

(a) an amino-flanking region comprising the sequence CPxxCxC.....xGxxVDCxxxGLx...xαPxxαPxDTTx,

(b) a leucine-rich repeat region comprising one or more repeats of the sequence xxxxFxxLxxLxLxxNxIxxL, and

(c) a carboxy-flanking region comprising the sequence P(W or F)xC(D or N)Cxα.....W(L or F)xxxxxxxxxxxxxx.....RCxx Pxxxxxxxαxxαxxxxxx...C(P or S).

The present invention is also directed to the following four amino-flank-LRR-carboxy-flank sequence elements of the SLIT protein.

(1) A first amino-flank-LRR-carboxy-flank sequence element of the SLIT protein (SEQ. I.D. NO. 3) wherein

(a) the amino-flanking region comprises the sequence: CPRVCSC TGLNVDCSHRGLT SVPRKISADVER,

(b) the leucine-rich region comprises the sequence:

LELQGNNLTVI
YETDFQRLTKLRMLQLTDNQIHTI
ERNSFQDLVSLERLDISNNVITTV
GRRVFKGAQSLRSLQLDNNQITCL
DEHAFKGLVELEILTLNNNNLTSL
PHNIFGGLGRLRALRLSDN

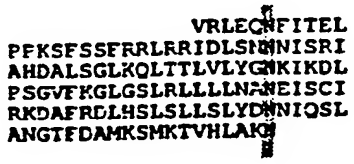
and

(c) the carboxy-flanking region comprises the sequence PFACD CHL SWLSRFLRSATRLAPYT RCQSPQLKGQNVADLHDQEFK CSGLTEHAPMECGAENS.

(2) A second amino-flank-LRR-carboxy-flank sequence element of the SLIT protein (SEQ. I.D. NO. 4) wherein

(a) the amino-flanking region comprises the sequence: CPHPCRC ADGIVDCREKSLT SVPVTLPDDTTD,

(b) the leucine-rich region comprises the sequence:



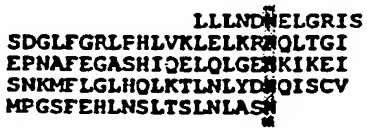
and

(c) the carboxy-flanking region comprises the sequence PFICNCNL RWLADYLHKIPIETSGARCESPKRMHRRRIESLREEKFK CSWGELRMKLSGECRMDSD.

(3) A third amino-flank-LRR-carboxy-flank sequence element of the SLIT protein (SEQ. I.D. NO. 5) wherein

(a) the amino-flanking region comprises the sequence: CPAMCHC EGTTVDCTGRGLK EIPRDIPLHTTE

(b) the leucine-rich repeat region comprises the sequence:



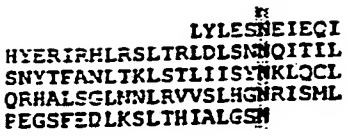
and

(c) the carboxy flanking region comprises the sequence: PFNCNCHL AWFAECVRKKSLNGGAA RCGAPSKVRDVQIKDLPH SEEK CSSENSEGCLGD GY.

(4) A fourth amino-flank-LRR-carboxy-flank sequence element of the SLIT protein (SEQ. I.D. NO. 6) wherein

(a) the amino-flanking region comprising the sequence CPPSCTC TGTVVACSRNQLK EIPRGIPAETSE,

(b) the leucine-rich repeat region comprises the sequence:



and

(c) the carboxy-flanking region comprises the sequence: PLYCDCGL KWFSDWIKLDYVEPGIA RCAEPEQMKDKLILSTPSSSFV CRGRVRNDILAKCNA.

The invention also relates to the alternate splice segment of the SLIT protein residing at the seventh epidermal growth factor (EGF) sequence element of the SLIT protein comprising the sequence GEGSTEPFTVT (SEQ. I.D. NO. 7).

The invention further concerns the carboxy terminal region of the SLIT protein (SEQ. I.D. NO. 9) residing after the seventh EGF.

Still further, the present invention is directed to combinations comprising one or more amino-flank-LRR-carboxy-flank sequence elements as defined above and one or more EFG-like repeat elements of the SLIT protein, provided that the combination does not include the naturally occurring configuration of the SLIT protein. The aforesaid combination can include the aforesaid alternative splice segment of the SLIT protein.

The present invention also encompasses antibodies to the SLIT protein or to the portions thereof encompassed by the present invention. Such antibodies are produced when the SLIT protein as described herein is introduced in an animal, e.g., a rabbit, mouse or rat, so as to raise antibodies in the animal and such antibodies are then withdrawn from the animal. The present invention is further directed to monoclonal antibodies to the SLIT protein or to the portions thereof encompassed by the present invention.

The invention also concerns diagnostics and therapeutics. Immunoassays are provided by the invention. In one such immunoassay a method for detecting the SLIT protein or a shed portion thereof in a bodily fluid from, for example, a human, is provided comprising contacting the bodily fluid with the antibodies to the SLIT protein described herein and detecting for the presence of the SLIT protein. Alternatively, a method of detecting autoimmune antibodies to the SLIT protein or a shed portion thereof in a bodily fluid from, for example, a human, is provided which comprises contacting the bodily fluid with the SLIT protein or portions thereof as defined herein and detecting for the presence of autoimmune antibodies to the SLIT protein.

The invention is also directed to detecting chromosomal rearrangements in the SLIT locus comprising hybridizing a nucleic acid (DNA or RNA) from a patient, e.g., a human patient, with a nucleic acid sequence from the SLIT (genomic) locus and detecting for the level of expression or an aberrant rearrangement.

The invention also relates to a pharmaceutical preparation for the treatment of a neurodegenerative disease, for treating traumatic injury to a neural tissue or for affecting the angiogenic process in a patient comprising a sterile (pharmaceutically acceptable) preparation comprising an effective amount of the SLIT protein as disclosed herein or to a portion thereof in admixture with a pharmaceutically acceptable carrier. The invention further includes the administration of such pharmaceutical preparation or a SLIT protein or a portion thereof, without a carrier, as disclosed herein or a portion thereof encompassed by the present invention in an effective amount to treat patients, e.g., humans, suffering from neurodegenerative disease or a

traumatic injury to a neural tissue or to affect the angiogenic process.

In addition, the invention is also directed to a class of multifunctional "TAGON" molecules which facilitate interactions between cell surface receptors involved in cell regulation and extracellular matrix molecules.

Thus the invention also concerns a protein, TAGON, that allows for the formation of a molecular bridge between axonally associated receptors and extracellular matrix molecules.

The invention also concerns a pharmaceutical preparation for the treatment of a neurodegenerative disease, for treating a traumatic injury to a neural tissue or for affecting the angiogenic process in a patient comprising a sterile preparation comprising an effective amount of a TAGON protein in admixture with a pharmaceutically acceptable carrier.

The present invention is also directed to a method for the treatment of a neurodegenerative disease, for treating tissue or for affecting the angiogenic process in a patient comprising administering to said patient an effective amount of a TAGON protein, either alone or in admixture with a pharmaceutically acceptable carrier.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1A schematically depicts the SLIT transcript. Fig. 1B is a restriction map of the genomic sequence containing the SLIT transcript.

Fig 2A schematically depicts the SLIT protein. Fig. 2B schematically depicts the elements of the SLIT protein.

Fig. 3 comprises twelve photographs depicting the SLIT message, protein, and promoter activation at three stages of embryogenesis by in situ hybridization, antibody staining and enhancer trap detection.

Fig. 4 comprise three photographs of an embryo undergoing dorsal closure stained with anti-SLIT antibodies.

Fig. 5 is a photograph depicting immunoelectron microscopic localization of SLIT in embryonic CNS to midline cells and axonal tracts.

Fig. 6 depicts immunoblots which show the secretion of SLIT from cultured cells.

Fig. 7 comprises eight photographs which show the pattern of expression of ß-galactosidase in MP2 cells and the midline neuroepitheium and its progeny compared in wild type and null mutant embryos.

Fig. 8 comprises six photographs which show that levels of SLIT expression correlate with disruptions of midline cells and axon pathways.

DETAILED DESCRIPTION OF THE DRAWINGS

Fig. 1. Transcription Unit and Molecular Characterization of SLIT P-element Enhancer Trap Alleles

In Fig. 1, the SLIT transcript (Fig. 1A) is shown aligned above the corresponding genomic sequence (Fig. 1B). Transcription is shown from left to right. Alternating light and dark shading patterns are used to represent the five EcoRl restriction fragments in the CDNA with the numbers above indicating their size in base pairs. Where known precisely,

the location of splice sites are shown by a connecting "v". Other exonic regions are shown as blocks aligned approximately with corresponding genomic sequence. The location of primers used to confirm the splice variation in the SLIT transcript and the resulting 33bp alternate segment are indicated by opposing horizontal arrows and a vertical bar, respectively. The location of the primer used to detect the P-element inserts is shown by a left pointing arrow near the 5' end of the transcript. Fig. 1B is a restriction map of the genomic sequence containing the SLIT transcription unit. Labeled triangles indicate the site of insertion of the enhancer trap construct in the various P-element SLIT alleles. Their nucleotide position relative to the consensus transcription initiation site is shown in parenthesis (B=BamHI; E=EcoRI; H=HindIII; S=SalI).

Fig. 2. Conservation of Flank-LRR-Flank Domains in Known Adhesive Proteins

Fig. 2A is a schematic representation of the SLIT protein. The putative signal sequence and amino and carboxy-terminal ends of the protein are indicated. The four consecutive Flank-LRR-Flank regions, the 7 EGF repeats and the 11 amino acid connecting segment, the result of differential splicing at the COOH-terminal of the 7th EGF repeat, are shown. Single LRRs have been shown to form ß-sheets in solution and, as depicted here, may form anti-parallel sheets (Krantz, D. E., and Zipursky, S. L. (1990). Drosophila chaoptin, a member of the leucine-rich repeat family, is a photoreceptor cell-specific adhesion molecule. EMBO J.9, 1969-1977). Tandem EGF-like repeats in other ECM proteins have been shown to be arranged in a rod-like conformation and are depicted in Fig. 2A as such (Engel, J. (1989). EGF-like domains in extracellular matrix proteins: localized signals for growth and differentiation. FEBS. 251, 1-7) with the individual EGF

repeats modeled after the solution structure of human EGF (Cooke, R. M. Wilkinson, A. J., Baron, M., Pastore, A., Tappin, M. J., Campbell, I. D., Gregory, H. and Sheard, B. (1987). The solution structure of human epidermal growth factor. Nature, 327, 339-341).

Fig. 3. Comparison of in situ, Antibody and Enhancer Trap Staining

The SLIT message, protein, and promoter activation are visualized at three stages of embryogenesis by in situ hybridization (A,D,G and J), antibody staining (B,E,H and K) and enhancer trap detection (C,F,I and L). The following stages during embryogenesis are shown; gastrulation in a dorsal view (A,B and C), germ band extended stage in a dorsal view (D,E and F) and nerve cord condensation, from both dorsal (G,H and I) and sagittal views (J,K and L). Staining can be demonstrated by all three methods in the midline neuroepithelium (arrow in D,E,F), midline glial cells (bold arrow in G,H,I,J,K) and cardioblast (open arrow in J,K,L), as well as in the walls of the gut and in a segmentally-reiterated pattern near the muscle attachment sites (thin arrow G,H,I). While no signal above background is detected from the lateral neuronal cell bodies, antibody staining (long thin arrow in H) is visible on the axonal projections from these neurons.

Fig. 4. Confocal Localization of the SLIT Protein to Cardioblasts and Muscle Attachment Sites

Fig. 4A depicts an optical, horizontal section of an embryo undergoing dorsal closure stained with anti-SLIT antibodies shows the SLIT protein to be localized on the surface of cardioblasts (opposing arrows) and at the muscle attachment sites to the body wall (long arrow).

Fig. 4B depicts a higher magnification view of the cardioblasts and shows that the highest concentration of the SLIT protein is localized to the regions of contact (long arrow) between opposing pairs of cardioblasts (apposing arrows) as they come together to form the lumen of the larval heart.

Fig. 4C is a sagittal view (dorsal side up) that shows the SLIT protein to be localized to the sites of muscle attached to the ectoderm (long arrows). Autofluorescence from the gut is also visible.

Fig. 5. Immunoelectron Microscopic Localization of SLIT in the Embryonic CNS to Midline cells and Axonal Tracts

Staining with anti-SLIT antibody in a frontal section through the plane of the longitudinal and commissural axonal tracts, detected by silver intensification of an HRP-conjugated secondary antibody. At the E.M. level labeling is both on the axons comprising the longitudinal connectives (lc), anterior (ac) and posterior (pc) commissures and on the cells lying between them including the processes of the midline glial cells (arrows). A light level frontal view of a similarly prepared dissected nerve cord shows strong axonal labeling with respect to the midline cells (see insert). No signal above background is seen on lateral neuronal cell bodies (N) either at the light or electron microscopic level. (scale bar = 5μm.)

Fig. 6. Secretion of SLIT from Cultured Cells

Fig. 6A depicts an immunoblot with anti-SLIT antibodies of the SLIT protein immunoprecipitated from embryos (Lane 1) and S2 culture cells Lane 2), shows a common protein species of approximately 200kD (arrow). This species is also

immunoprecipitated from S2 cell line conditioned media (Lane 3) indicating that the SLIT protein can be exported from the cells in which it is produced. Lane 4 shows by immunobloting that the 200kD SLIT protein species can also be detected in the matrix materials deposited by the S2 cells in culture. Predominant band seen in immunoprecipitations is immunoglobulin heavy chain (indicated by an H).

In Fig. 6B the media in which S^{35} metabolically-labeled S2 cells had been cultured was immunoprecipitated with anti-SLIT antibodies, separated by SDS-page, and detected by autoradiography. Consistent with the immunoblotting results, a major 200 kD species is detected (arrow). Tick marks indicate position of 100 kD and 220 kD molecular weight size standards.

Fig. 7. Null Mutant Embryos Exhibit Disruptions in Midline Cells

The pattern of expression of ß-galactosidase in the MP2 cells (A,B) and the midline neuroepithelium and its progeny (C-H) is compared in wild type and null mutant embryos. Anterior is toward the left.

(A,B): A dorsal view shows the MP2 cells (arrows) well separated by cells of the midline neuroepithelium at the extended germband stage in wild-type embryos (A) but closer together in a SLIT mutant background (B), indicating an early disruption along the midline.

(C,D): The midline neuroepithelium at the germband extended stage (arrow in C) and its midline progeny (E,G) are clearly labeled in wild type embyros. In comparison, following germband extension in slit mutant embryos there is either no midline neuroepithelial labeling, or low levels of labeling slightly later (arrow in D).

(E,F): A sagittal view during nervecord condensation shows the bulk of the midline cells of each neuromere clearly expressing

ß-galactosidase in the wild-type embryo (arrow in E). However, in slit mutant embryos, the expressing cells are reduced in number and displaced to the ventral edge of the nerve cord (arrow in F).
(G,H): A dorsal view of a similarly staged wild type (G) and SLIT mutant (H) embryo. In the wild type the midline cells can be seen in the space separating adjacent neuromeres within a segment. In SLIT mutant embryos, expressing cells can be seen to lie irregularly shifted laterally as well as ventrally (arrow).

Fig. 8. Levels of SLIT Expression Correlate with Disruptions of Midline Cells and Axon Pathways

The major axonal pathways are labeled with anti-HRP antibodies (A,C,E) (Jan, L. Y., and Jan, Y. N. (1982). Antibodies to horseradish peroxidase as specific neuronal markers in Drosophila and grasshopper embryos. Proc. Natl. Acad. Sci. USA 79, 2700-2704) and compared to the staining pattern seen with antibodies against the SLIT protein (B,D,F). In these horizonal views anterior is toward the left.
(A,B): In wild type embyros the ladder-like arrangement formed by the commissural and longitudinal axonal tracts is visible. Staining with antibodies against the SLIT protein (B) shows labeling of the midline glial cells (thick, mid-sized arrow) as well as axonal staining (short arrow).
(C,D): Anti-HRP stained null mutant embryos (C) exhibit a single centrally located longitudinal nerve bundle along the length the CNS. No detectable SLIT staining is seen (D). The lateral neuronal bodies are shifted inward toward the center, filling the space normally occupied by the midline cells. An overall reduction in the width of the nervecord is also observed (double-ended arrow).

(E,F): $slit^{E158}$ mutants exhibit an intermediate phenotype characterized by a partial collapse of the axonal scaffold. Relatively weak SLIT staining is visible along the length of the axonal bundles (F). Segments with the highest levels of SLIT staining (arrow), have more midline cells and a less severe collapse of the longitudinal connectives (short arrow) in comparison to segments with lower expression levels (long arrow). Segments with reduced levels of slit expression exhibit nervecord compression and a concomitant fusion of the axon tracts (long arrow).

BRIEF DESCRIPTION OF THE SEQUENCES

SEQ. I.D. NO. 1

The SLIT Nucleotide Sequence Codes for a Putative Extracellular Protein with Both Flank-LRR-Flank and EGF domains

The cDNA sequence containing the slit coding region is shown as SEQ. I.D. NO. 1. The coding domain is characterized by the presence of a putative signal sequence and four distinct blocks of leucine-rich repeats followed by two regions containing epidermal growth factor repeats. The location of the predicted signal sequence cleavage site is indicated. There are 13 potential N-linked glycosylation and two consensus sequences for ß-hydroxylation (Rees, D. J. G., Jones I. M., Handford, P.A., Walter, S. J., Esnouf, M. P., Smith, K. J., and Brownlee, G. G. (1988)). The role of ß-hydroxyaspartate and adjacent carboxylate residues in the first EGF domain of human factor IX. EMBO J. 7,2053-2061) in the third and fifth EGF repeats. The 33bp alternatively spliced segment in the slit transcript, and the 11 amino acids which it encodes are shown.

SEQ. I.D. NO. 2
Amino acid sequence of the entire SLIT protein, including four Amino-flank-LRR-Carboxy-flank domains, 6 tandem EGF-like repeats, an intervening region, the 7th EGF-like repeat, an alternative splice segment, and a carboxy terminal region.

SEQ. I.D. NO. 3
Amino acid sequence of the first Amino-flank-LRR-Carboxy-flank domain of SLIT protein.

SEQ. I.D. NO. 4
Amino acid sequence of the second Amino-flank-LRR-Carboxy-flank domain of SLIT protein.

SEQ. I.D. NO. 5
Amino acid sequence of the third Amino-flank-LRR-Carboxy-flank domain of SLIT protein.

SEQ. I.D. NO. 6
Amino acid sequence of the fourth Amino-flank-LRR-Carboxy-flank domain of SLIT protein.

SEQ. I.D. NO. 7
Eleven amino acid alternative splice segment.

SEQ. I.D. NO. 8
Concensus amino acid sequence for an Amino-flank-LRR-Carboxy-flank domain.

SEQ. I.D. NO. 9
Carboxy-terminal region of the SLIT protein.

DETAILED DESCRIPTION OF THE INVENTION

The present invention concerns the full structure of the SLIT protein, sequence elements thereof and the design of diagnostic and therapeutic reagents based on the elucidation of their role in biological systems.

Applicants found that, in addition to containing EGF homologous domains, the SLIT protein also has four regions bearing homology to the leucine-rich repeats (LRRs) found in a family of proteins involved in protein-protein interactions (Titani, K., Takio, K., Handa, M., and Ruggeri, Z. M. (1987). Amino acid sequence of the von Willebrand factor-binding domain of platelet membrane glycoprotein Ib. Proc Natl. Acad, Sci, USA 84, 5610-5614; Schneider, R., Schneider-Scherzer, E., Thurnher, M., Auer, B., and Schweiger, M. (1988). The primary structure of human ribonuclease/angiogenin inhibitor (RAI) discloses a novel highly diversified protein superfamily with a common repetitive module. EMBO. J. 7, 4151-4156; McFarland, K.C., Sprengel, R., Phillips. H. S., Kohler, M., Rosemblit, N., Nikolics, K., Segaloff, D. L., and Seeberg, P.H. (1989). Lutropin-Choriogonadotropin Receptor: An unusual member of the G protein-coupled receptor family. Science 245, 494-499; Field, J., Xu, H-P., Michaeli, T., Ballester, R., Sass, P., Wigler, M., and Colicelli, J. (1990) Mutations of the adenylyl cyclase gene that block RAS function in Saccharomyces

cervisiae. Science 247, 464-467; Krantz et al (1990) EMBO, J. 9, 1969-1977).

In addition, it is demonstrated herein that sequences flanking the LRRs of SLIT exhibit homology to sequences in corresponding positions in some of the other LRR-containing proteins. It is also demonstrated herein that SLIT is necessary for the normal development of the midline of the CNS, including in particular the midline glial cells, and for the concomitant formation of the commissural axon pathways. Furthermore, this process is dependent on the level of SLIT protein expression. Data is provided herein indicating that the SLIT protein is excreted from the midline glial cells where it is synthesized, and is eventually associated with the surfaces of the axons that traverse them. In addition, the SLIT protein is tightly localized to the muscle attachment sites and to the sites of contact between adjacent pairs of cardioblasts as they coalesce to form the lumen of the larval heart. The implications of the structure and distribution of the SLIT protein in development are discussed in detail hereinbelow.

Molecular Characterization of the SLIT Transcript and P-element Alleles

The isolation and partial characterization of SLIT EGF-homologous sequences and corresponding cDNA clones was

described previously (Rothberg et al. 1988 supra). Applicants have extended this molecular analysis to include the entire SLIT coding sequence, its genomic organization, characterization of a splicing variant, and the molecular basis of four P-element induced mutations. The SLIT embryonic transcript was estimated to be approximately 9kb by Northern analysis. Using both conventional hybridization screening procedures and methods employing the polymerase chain reaction (PCR), applicants obtained cDNA clones representing 8.6kb of this sequence. Sequencing of genomic DNA indicates a consensus Drosophila transcriptional initiation sequence (Hultmark, D., Klemenz, R., and Gehring, W. J. (1986). Translational and transcriptional control elements in the untranslated leader of the heat-shock gene hsp22. Cell 44,429-438, 1986) 53bp upstream of applicants' longest cDNA.

Fig. 1 shows the SLIT transcript aligned with a restriction map of the corresponding genomic regions. The known intron/exon boundaries are indicated in Fig. 1A and were determined by a comparison of the cDNA sequence with known genomic sequence (Rothberg et al., 1988 supra). The SLIT cDNA sequence spans an approximately 20kb genomic region and contains a single 4440 bp open reading frame (ORF). The nucleotide and deduced amino acid sequences of the ORF are shown in SEQ. I.D. NO. 1.

The SLIT coding sequence (Gibskov, M., Devereux, J., and Burgess, R.B. (1984). The codon preference plot: graphic analysis of protein coding sequences and prediction of gene expression. Nucl. Acid. Res. 12, 539-549) starts with a translational start site consistent with the Drosophila consensus (Cavener, D.R. (1987) Comparison of the consensus sequence flanking translational start sites in Drosophila and vertebrates. Nucl. Acid. Res. 15, 1353-1361).

Restriction mapping and sequence analysis of SLIT cDNA clones revealed two classes of transcript differing by 33 nucleotides. The location of this sequence variation is shown in SEQ. I.D. NO. 1. The presence of a minor sequence variation prompted a more careful analysis of slit cDNA clones in order to detect whether other transcript variants existed that might not have been detected by Northern analysis. Utilizing a cDNA screening procedure based on the PCR, the only detectable size variation was confined to the same region as in the original variant. A comparison of the genomic and cDNA sequences demonstrates that the 33 nucleotide size variation is the result of alternate RNA splicing. The two species of SLIT cDNA differ in the location of a donor (5') splice site, while the acceptor (3') site is identical.

The molecular characterization was been extended to include the determination of the site of P-element insertion

in four SLIT alleles $slit^{F81}$, $slit^{F119}$, $slit^{E158}$ and $slit^{175}$, which were recovered during a P-element based enhancer trap screen (Bier E, et al., (1989)). Searching for pattern and mutation in the Drosophila gerome with a P-lacZ vector. Genes & Dev. 3, 1273-1287; Bellen, H. J., O'Kane C. J. Wilson, C., Grossniklaus, U., Pearson, R. K. and Gehring, W. Y. (1989) P-element-mediated enhancer detection: a versatile method to study development in Drosophilia, Gen. & Dev. 3, 1273-1287). Genomic DNA from each line was employed in the PCR using primers designed to detect P-element insertions in regions 5' of the SLIT coding sequence. By direct sequencing of the PCR products, these lines were shown to contain insertions upstream of both the SLIT consensus transcription initiation sequence and ORF (see Fig. 1B) confirming their initial characterization as SLIT alleles and suggesting their utility in the characterization of SLIT expression.

SLIT Codes for Flank-LRR-Flank and EGF Domains

The SLIT transcripts potentially encode two proteins of 1469 and 1480 amino acids, with molecular weights of approximately 166kD. The predicted initiating methionine is followed by an amino acid sequence containing structural regions characteristic of a secretory signal sequence (SEQ I.D. NO. 1). However, hydropathy plots do not predict a transmembrane domain (data not shown). An examination of the

SLIT coding domain reveals that the majority of the protein is composed of two repeated motifs: the 24 amino acid leucine-rich repeat (LRR) and the 40 amino acid EGF repeat (SEQ. I.D. NO. 1). Fig. 2A shows schematically the positions of these repeats and indicates a higher level of organization among the LRRs. The LRRs are arranged in four groups, each composed of four or five LRRs surrounded by conserved amino- and carboxy-flanking regions (Fig. 2B) (SEQ. I.D. NOS. 3,4, 5 and 6). The presence of both the LRRs and EGF-like repeats within a single protein make SLIT unusual; this combination is not found in any other proteins in the NBRF databank. The absence of any potential transmembrane domains in a sequence having a typical signal sequence and two known extracellular-associated motifs suggests that the SLIT locus encodes a secreted extracellular protein.

The LRR motif is found in a variety of vertebrate and invertebrate proteins involved in protein-protein interactions (Table 1).

TABLE 1

Table 1. Leucine-rich Repeat Containing Proteins

Proteins		Arrangment	Function	Reference
Glycoprotein Ibα		LRR-Flank	Receptor/Adhesion	Titani et al., 1987; Lopez et al., 1987
Glycoprotein Ibß		Flank-LRR-Flank	Receptor/Adhesion	Lopez et al., 1988
Glycoprotein IX		Flank-LRR-Flank	Receptor/Adhesion	Hickey et al., 1989
Lutropin-Choriogonadotropin receptor		LRR	Receptor	McFarland et al., 1989
Collagen-binding 59 kd protein (fibromodulin)		Flank-LRR	ECM binding	Oldberg et al., 1989
Small interstitial proteoglycan PG-S1 (Biglycan)		Flank-LRR	ECM binding	Fisher et al., 1989
Small interstitial proteoglycan PG-S2 (Decorin, PG-40)		Flank-LRR	ECM binding	Krusius et al., 1986; Day et al., 1987
Adenylate cyclase[a]		LRR	Protein-Protein	Kataoka et al., 1985; Field et al., 1990
Ribonuclease/angiogenin inhibitor[a]		LRR	Protein-Protein	Schneider et al., 1988
Chaoptin		LRR	Homotypic Adhesion	Reinke et al., 1988; Krantz et al., 1990
Leucine-rich α_2 glycoprotein		LRR	??	Takahashi et al., 1985
Oligodendrocyte-myelin Glycoprotein		Flank-LRR	Adhesion?	Mikol et al., 1990
Toll	2x	LRR-Flank	Dorsal-ventral polarity[b]	Hashimoto et al., 1988
slit	4x	Flank-LRR-Flank	Morphogenesis [b]	This invention

[a] Intracellular proteins, all other are extracellular or cell surface proteins.

[b] While the role of these proteins in Drosophila development is known, it is not known how their function is mediated.

References Listed in Table 1

Lopez, J. A. Chung, D. W. Fujikawa, K., Hagen, F. S., Davie, E. W., and Roth, G. J., (1988). The α and ß chains of human platelet glycoprotein IB are both transmembrane proteins containing a leucine-rich amino acid sequence. Proc. Natl. Acad. Sci. USA 85, 2135-2139.

Hickey, M. J., Williams, S. A., and Roth, G. J. (1989). Human platelet glycoprotein IX: An adhesive prototype of leucine-rich glycoproteins with flank-center-flank structures. Proc. Natl. Acad. Sci. USA 86,6773-6777.

Oldberg, A., Antonsson, P., Lindblom, K., and Heinegard, D. (1989). A collagen-binding 59-kd protein (fibromodulin) is structurally related to the small interstitial proteoglycan PG-S1 (decorin). EMBO J. 8, 2601-2604.

Fisher, L. W., Termine, J. D., and Young, M. F. (1989). Deduced protein sequence of bone small proteoglycan (Biglycan) shows homology with proteoglycan II (Decorin) and several nonconnective tissue proteins in a variety of species. J. Biol. Chem. 264, 4571-4576.

Krusius, T., and Ruoslahti, E. (1986). Primary structure of an extracellular matrix proteoglycan core protein deduced from cloned cDNA. Proc. Natl. Acad. Sci USA 83,7683-7687.

Kataoka, T., Broek, D., and Wigler, M. (1985). DNA sequence and characterization of the S. Cerevisie gene encoding adenylate cyclase. Cell 43, 493-505.

Reinke, R., Krantz, D. E. Yen, D., and Dipursky, S. L. (1988). Chaoptin, a cell surface glycoprotein required for Drosophila photoreceptor cell morphogenesis, contains a repeat motif found in yeast and human. Cell 52,291-301.

Takahashi, N., Takahashi, Y. and Putnam, F. W. (1985). Periodicity of leucine and tandem repetition of a 24-amino acid segment in the primary structure of leucine-rich a2-glycoprotein of human serum. Proc. Natl. Acad. Sci. USA 84,4767-4771.

Mikol, D.D., Gulcher, J. R. and Stefansson, K. (1990). The Oligodendrocyte-Myelin Glycoprotein Belongs to a Distinct Family of Proteins and Contains the HNK-1 Carbohydrate. J. Cell Bio. 110,471-479.

Together with their surrounding sequences, the tandem arrays of LRRs in SLIT form a Flank-LRR-Flank structure, part of which was previously noted in some of the other LRR-containing proteins (Hickey et al., 1989 supra). However, in this application, applicants extend both the amino-terminal LRR flanking sequence and the carboxy-terminal flanking sequences to include invariant cysteines, arginines, prolines, and other conserved residues (consensus in SEQ. I.D. NO. 8). A comparison of other LLR-containing proteins with SLIT reveals that a subset have homology to SLIT extending to either one or both of the conserved flanking regions as defined herein (Table 1; SEQ. I.D. NO. 8). This similarity is found in the oligodendrocyte-myelin glycoprotein (OMgp) of humans, the Toll gene of Drosophila melanogaster and among two sets of structurally related vertebrate proteins involved in adhesive events. OMgp is believed to mediate the adhesion of oligodendrocytes to either other glial cells or axons (Mikol, et al., 1990 supra) and contains the amino-flanking region and 7 LRRs. Toll, a transmembrane protein, is required for dorsal-ventral pattern formation (Hashimoto, C., Hudson, K. L., and Anderson, K.V. (1988). The Toll gene of Drosophila, required for dorsal-ventral embryonic polarity, appears to encode a transmembrane protein. Cell 52,269-279) and has an

extracellular domain characterized by the presence of two LRR regions with SLIT homologous carboxy-flanking sequences.

The first set of vertebrate proteins with slit homology in their flanking regions comprise the von Willebrand factor receptor (Titani et al., 1987, supra; Lopez et al., 1988, supra; Hickey et al., 1989, supra). The similarities between SLIT and two members of this protein complex, GPIX and GPIbß, include the full Flank-LRR-Flank motif, albeit with a single LRR. The third member of this complex GPIbα, however, contains a tandem array of LRRs and a conserved carboxy-flanking region without a conserved amino-flanking region. Extensive similarity between SLIT and a second group of vertebrate proteins is apparent in their LRR and amino-flanking regions. This group consists of the ECM proteoglycans decorin (Day, A. A., McQuillan, C. I., Termine, J. D., Young, M. R. (1987). Molecular cloning and sequence analysis of the cDNA for small proteoglycan II of bovine bone. Biochem. J. 248, 801-805; Krusius and Ruoslahti, 1986, supra) and biglycan (Fisher et al., 1989, supra) and fibromodulin (Oldberg, 1989, supra). These proteins have overall homology to one another and define a family of extracellular proteins with conserved amino-flanking regions and 10 consecutive LRRs (Oldberg et al., 1989, supra).

All the proteins exhibiting homology to SLIT in their LRR flanking regions have either been shown, or are believed, to participate in extracellular protein-protein interactions. Moreover, SLIT contains 7 copies of the EGF motif (Fig. 2A), which also has been shown to participate in extracellular protein-protein interactions (Rothberg et al., 1988 supra). The last EGF repeat is of special interest because the alternate mRNA splicing noted earlier potentially results in the insertion or removal of 11 unique amino acids at the COOH terminal of this repeat (see SEQ. I.D. NO. 7) (Fig. 2A).

SLIT is Exported From Glial Cells and Distributed Along Axon Tracts

It has been shown previously that SLIT transcript and protein could be detected at the highest levels in the midline glial cells (Rothberg et al., 1988 supra). However, despite the presence of the SLIT protein on the axons in the embryonic commissural and longitudinal axon pathways, applicants failed to detect any transcript or protein in the cell bodies of these neutrons. This raised the possibility that the SLIT protein, which is synthesized in and presumably secreted by the midline cells, can become associated with axons. Here this question is further explored in whole-mount embryo preparations by comparing the sites of SLIT expression, as assayed by in situ hybridization and the detection of ß-galactosidase in SLIT enhancer trap lines, with the subsequent localization of the protein as assayed by antibody staining (summarized in Fig. 3).

All four enhancer trap alleles ($slit^{F81}$, $slit^{F119}$, $slit^{E158}$ and $slit^{175}$) express ß-galactosidase within the ventral midline to varying levels. The location of the P-element constructs 5' of the SLIT coding domain, the resulting mutant phenotypes and especially their expression patterns are all consistent with their being under the transcriptional control of SLIT regulatory elements. A summary of the embryonic localization of the SLIT mRNA and protein, and the ß-galactosidase expression of $slit^{E158}$ is shown in Fig. 3. The expression of ß-galactosidase from the enhancer trap construct in $slit^{E158}$ shows excellent overall agreement with mRNA localization data at all embryonic stages (compare Figs. 3A, D, G and J with 3C, F, I and L). Each method reveals a nearly identical expression pattern starting at gastrulation (Fig. 3A, B and C). At germband extension, all of the midline mesectodermal cells (Crews, S.T., Thomas, J. B., and Goodman, C. S. (1988)

The Drosophila single-minded gene encodes a nuclear protein with sequence similarity to the per gene product. Cell 52, 143-151; Thomas, J. B., Crews, S. T. and Goodman, C. S. (1988). Molecular genetics of the single-minded locus: a gene involved in the development of the Drosophila nervous system. Cell 52, 133-141) show the highest level of slit expression (Figs. 3D, E and F). During germband retraction and nerve cord shortening, expression is most restricted to the six midline glial cells which are derivatives of the midline neuroepithelium (Figs. 3G, H, and I). Localized expression is also evident in the cardioblasts (Figs. 3J, K and L) during dorsal closure. Figs. 4A and B show that the SLIT protein is most highly localized to the points of contact between opposing pairs of cardioblasts as they coalesce to form the dorsal vessel (presumptive larval heart). All three methods also reveal expression in the walls of the gut (Figs. 3J, K and L) and in a segmentally-reiterated pattern near the muscle attachment sites in the ectoderm (apodemes; Figs. 3G, H and I). Precise protein localization to the sites where the muscles are attached to the apodemes is seen by confocal microscopy (Figs. 4A and C).

In situ hybridization (Figs. 3D, G and J) and the expression from the enhancer trap lines (Figs. 3F, I and L) both support the observation that initially all of the midline cells, and subsequently primarily the six midline glia, are producing SLIT while lateral neurons are not. However, antibody labeling is seen strongly in the midline glia (Fig. 3E, H) and on the commissural and longitudinal axon tracts (Fig. 3E, H and K), while it is absent from lateral neuronal cell bodies, which supply the bulk of the axons to these bundles. These results suggest that the antibody labeling along the commissural and longitudinal axon tracts is due to the distribution of SLIT protein exported from the midline

glial cells. The protein is also absent from the peripheral nerve roots and peripheral axon tracts.

Immunoelectron microscopy was used to determine the subcellular localization of the SLIT protein in the ventral nerve cord. Dissected embryonic nerve cords demonstrate staining on the midline cells as well as on the commissural and longitudinal nerve bundles. Light and electron micrographs of a similarly prepared sample are shown in Fig. 5. While all the derivatives of the neuroepithelium initially express SLIT, during nerve cord condensation and axonal outgrowth this expression becomes restricted to the midline glial cells. The midline glial cells surround the developing commissural axons and growth cones have been shown to track along their surface (see Jacobs and Goodman, 1989, J. Neurosci., 9, 2402-2411). Antibody staining can be seen both on the surface of the midline glial cells where they abut growing axons and on the axons themselves. No detectable variation in the amount of SLIT staining among subsets of axons or fascicles is detected.

Applicants were able to detect SLIT along the length of the axonal projections in the commissural and longitudinal axon tracts though we are unable to detect any signal above background from the lateral neuronal cell bodies supplying these axonal (Fig. 5). Immunoelectron microscopy demonstrated the extracellular localization of the SLIT protein and supports the expression data indicating that the SLIT protein on the axon tracts is not produced by the neurons whose axons comprise them. Thus, it appears that the axonally distributed SLIT protein is first secreted from the midline glial cells and then becomes associated with these axons as they traverse the midline.

To obtain direct biochemical evidence that SLIT is exported from the cells in which it is produced, applicants investigated SLIT expression in Drosophila tissue culture cell lines. Schneider line S2 was found to normally express the SLIT protein, and it can be seen on the surface of a subset of the cells by immunofluorescence. Immunoblotting of immunoprecipitated protein extracts from Drosophila embryos and S2 cell lines revealed a singe 200kD molecular weight band (Fig. 6A, lanes 1 & 2). This size is consistent with expectations of a glycosylated form of the predicted SLIT protein. Conditioned Schneider cell media also was found to contain a similar 200kD molecular weight species (Fig. 6A, lane 3) in addition to two other species which may represent differences in glycosylation. The presence of the SLIT protein in the culture media was confirmed by immunoprecipitations of the same molecular weight species from media in which S^{35} metabolically-labeled S2 cells had been growing (Fig. 6B). These experiments further support the suggestion that SLIT is an excreted protein. Additionally, immunoblotting of the matrix materials deposited in culture by S2 cells showed the SLIT protein to be enriched in this fraction (Fig. 6A, lane 4), consistent with the hypothesis that SLIT functions as an extracellular matrix molecule.

SLIT Mutants Exhibit Disruptions in Midline Cells and Commissural Axon Pathways

An analysis of SLIT null mutant embryos reveals the collapse of the normal scaffold of commissural and longitudinal axons. However, the SLIT protein is detectable in the midline neuroepithelial cells well before the time of axonal outgrowth (Rothberg, 1988 supra). This raised the possibility that the SLIT protein influences the differentiation of midline cells from the neuroepithelium and that the observed collapse of the axonal scaffold is the

result of an earlier developmental abnormality. In order to examine the development of the midline before axon outgrowth, applicants followed the fate of the MP2 cells (an identified neuronal precursor cell that normally develops in the most medial row of neuroblasts in the lateral neuroepithelium) as well as the midline neuroepithelium and its progeny in both wild-type and mutant embryos.

In wild-type embryos at the germband-extended stage the MP2 cells are separated by the midline neuroepithelium (Fig. 7A), whereas in SLIT embryos these cells appear closer together (Fig. 7B). In addition, cell autonomous markers (lines 8-7 & 242) for some of the midline neuroepithelial cells and their progeny (Fig. 7C, E, G) are either absent or ectopically expressed before (Fig. 7D) and during axonal outgrowth (Fig. 7F, H). For example, in SLIT mutant embryos, some of these cells appear absent and others come to lie in an abnormal position along the ventral surface of the nervecord (Fig. 7F, H). These results clearly show a perturbation in the development of the midline neuroepithelial cells as early as the extended-germband stage. This disruption further leads to a disruption of their progeny, including the midline glial cells, resulting in a lateral compression of the nerve cord (confirmed by histological analysis). Given the disruption in the development of the midline of the CNS, the ensuing collapse of the axonal scaffold is not unexpected (a similar phenotype of the stimulant; Crews et al., 1988, supra; Thomas et al., 1988, supra).

Mutations caused by the insertion of the enhancer-trap P-element allow for a further exploration of the relationship between the level of SLIT expression and the extent of the nerve cord defect. In the wild-type embryo, as observed with antibodies specific to neuronal membranes, commissural and longitudinal axon pathways appear to form a regular ladder-

like structure (Fig. 8A). A wild-type embryo stained with anti-SLIT antibodies also shows labeling of the CNS axon pathways, as well as prominent staining of the midline glial cells (Fig. 8B). Embryos homozygous for slit^{IG107} do not have any detectable SLIT expression either in the midline cells or on the axonal bundles (Fig. 8D). Thus null allele is embryonic lethal; mutant embryos exhibit a lateral compression of the nerve cord (Fig. 8D), and a single fused longitudinal axon tract (Figure 8C).

As judged by antibody staining intensity in whole-mount embryo preparations, all four enhancer trap SLIT alleles show reduced levels of SLIT expression in the homozygous state at 18°C and exhibit an intermediate phenotype. Since the P-element construct resides upstream of SLIT coding sequences, it is reasonable to assume that it is not the disruption of the SLIT protein per se that is responsible for the observed mutant phenotypes, but rather a reduction in the level of SLIT expression. These mutations are embryonic and larval lethals and in contrast to the null allele slit^{IG107}, exhibit only partial compression of the midline and a concomitant partial collapse of the axonal scaffold (Fig. 8E and F). Variable levels of SLIT expression in the midline cells, often at lower levels and in a more diffuse pattern were noted compared to wild type. This variability is seen both between individual embryos and between segments in the same embryo (Fig. 8F). The segments with the lowest levels of expression exhibit the least differentiation their midline cells, including their midline glia, and show the greatest degree of collapse of both the ventral nerve cord and the axon tracts (Fig. 8F). Segments exhibiting higher levels of expression appear at a gross level to have nearly normal midline glial cells, commissures, and longitudinal axon tracts (Fig. 8F).

As mentioned herein, it is demonstrated herein that the SLIT locus, whose mutant phenotypes indicate that it plays a major role in the development of the specialized midline glial cells and the commissural axon tracts that traverse them, encodes a unique extracellular protein containing two structural motifs associated with adhesive interactions. The SLIT protein has four regions containing tandem arrays of a 24 amino-acid leucine-rich repeat (LRR) with conserved flanking sequences (Flank-LRR-Flank) and two regions with epidermal growth factor (EGF)-like repeats. Although the LRR and EGF motifs are not found together in any other proteins in the NBRF data bank, each has been found in conjunction with other sequence motifs, often forming a distinct region of a larger protein involved in protein-protein interactions. As part of larger proteins, each of these motifs has been shown to directly contribute to these interactions.

The LRRs in SLIT are similar to those that were first identified in human leucine-rich α2-glycoprotein and later in a variety of vertebrate and invertebrate proteins involved in protein-protein interactions, both inside and outside the cell (Table 1). In the extracellular environment, the LRRs have been found in conjunction with a variety of conserved protein motifs (McFarland et al., 1989 supra; Mikol et al., 1990 supra). Of greatest interest, however, is the fact that the LRRs in extracellular proteins are often found accompanied by either one or both of the conserved amino- and carboxy-flanking regions identified in the slit protein (see Table 1). In all of the cases where the LRR are accompanied by these flanking regions the proteins have either been shown, or are believed, to participate in extracellular adhesive interactions. While the significance of the individual flanking regions in these interactions is not yet known, a functional role for at least the carboxyl-flanking sequence has been demonstrated in vivo: mutations in the cysteines of

this region in the Drosophila Toll protein confer a dominant phenotype.

In addition to Toll and the oligodendrocyte-Myelin glycoprotein, two distinct families of adhesive proteins have SLIT homology extending to the LRR flanking sequences. The first includes a set of functionally related interstitial proteoglycans known to bind directly to ECM proteins: biglycan, fibromodulin and decorin. Biglycan binds laminin and fibronectin, while fibromodulin and decorin bind collagen and fibronectin and have a regulatory effect on collagen fibril formation (Vogel K. G., Paulsson M., and Heinegard, D. (1984). Specific inhibition of type I and type II collagen fibrillogenesis by the small proteoglycan of tendon. Biochem. J. 223, 587-597; Hedbom, E., and Heinegard, D. (1989). Interactions of 59-kDa connective tissue matrix protein with collagen I and collagen II. J. Biol. Chem. 264, 6898-6905; Oldberg et al(1989) supra; Schmidt, G., Robeneck. H., Harrach, B., Glossl, J., Nolte, V., Hormann, H., Richter, H., and Kresse, H. (1987). Interactions of small dermatan sulfate proteoglycan from fibroblasts with fibronectin. J. Cel. Biol. 104, 1683-1691). The second set comprises the proteins of the glycoprotein Ib-IX (GPIb-IX) complex, which together function as a receptor for the von Willebrand factor (vWF) and thrombin and are responsible for vWF-dependent platelet to blood vessel adhesion. In this complex, the LRR-containing region of the GPIbα chain binds one of a set of three repeated 200 amino acid sequences termed A domains in vWF (Titani et al, 1987 supra; Mohri H., (1988) Structure of the von Willebrand Factor Domain Interacting with Glycoprotein Ib. J. Biol. Chem., 17901-17904). In addition to demonstrating the role of the LRR motif in protein-protein interactions this homology also raises the possibility that similar regions in SLIT might bind to proteins containing repeats homologous to the A domains of vWF. In vertebrates, these proteins include both ECM

molecules and integrins (Larson, R. S., Corbi, A. L., Berman L., and Springer, T. (1989). Primary structure of the leukocyte function-associated molecule-1 α subunit: an integrin with an embedded domain defining a protein superfamily. J. Cell Bio., 108, 703-712).

The conservation of the amino-terminal sequences flanking a LRR region in a family of proteins that participate in direct adhesion to ECM components suggests that this structure may play a similar role in SLIT. Alternatively the conversation of the entire Flank-LRR-Flank motif in SLIT and the GPlb-IX complex offers the intriguing possibility that SLIT'S interactions with the ECM, like those of the vWF and thrombin receptor, could be mediated by additional factors.

In comparing the various proteins known to contain the EGF-like motif, it is clear that this sequence is always found in an extracellular environment and in many instances these sequences have either been implicated, or shown, to function directly in protein-protein interactions. (Apella, E., weber, I.T., and Blasi, F. (1988). Structure and function of epidermal growth factor-like regions in proteins. FEBS. 231, 1-4).

In addition, these repeats are found in conjunction with a variety of other structural and catalytic domains in molecules involved in blood coagulation (Furie, B., and Furie, B. C. (1988). The molecular basis of blood coagulation. Cell 53, 505-518) and in adhesive ECM glycoproteins (Engel, 1989, FEBS, 251, 1-7)). Tandem arrays of EGF-like repeats comprise the majority of the extracellular domains of the cell surface proteins Notch (Wharton, K. A., Johansen, K. M., Xu, T., and Artavanis-Tsakonas, S. (1985). Nucleotide sequence from the neurogenic locus Notch implies a gene product that shares homology with proteins containing EGF-like repeats.

Cell 43, 567-581) and Delta (Vassin, H., Bremer, K. A., Knust, E., and Campos-Ortega, J. (1987). The neurogenic gene Delta of Drosophila Melanogaster is expressed in neurogenic territories and encodes and putative transmembrane protein with EGF-like repeats. EMBO J. 6, 3431-3440). 1987; Kopczynski et al., 1988) and have been implicated in Ca++ - dependent heterotypic adhesive interactions between the two proteins as well as in homotypic interactions in the Delta protein (Fehon et al., 1990, Cell, 61, 523-534).

The EGF-like repeats in SLIT are arranged in two groups in a fashion similar to the arrangement found in cell surface and extracellular adhesive proteins and in EGF-like ligands, respectively (Apella et al., 1988 supra; Lander, A. D. (1989). Understanding the molecules of neural cell contacts: emerging patterns of structure and function. TINS. 12, 189-195). An additional similarity between the EGF-like repeats in SLIT, Delta and Notch is a conserved recognition site for a post-translational modification involved in Ca++ binding (Rees et al., 1988, EMBO J. 2053-2061) and a consensus sequence implication in Ca++ dependent protein-protein interactions (Handford, P.A., Baron, M., Mayhew, M., Willis, A., Beesley, T., Brownlee, G. G., and Campbell, I.D. (1990). The first EGF-like domain from human factor IX contains a high-affinity calcium binding site. EMBO J. 9, 475-480).

By these criteria the 3rd and 5th EGF-repeats of SLIT are potential candidates for ß-hydroxylation and may participate in Ca++ dependent interactions. The 7th and last EGF domain in SLIT is separated from the tandemly arranged EGF-repeats by 202 amino acids.

Export and Cell Binding

Using both whole-mount in situ hybridization and SLIT enhancer trap alleles, applicants were able to demonstrate

that SLIT is produced in the developing midline neuroepithelium, as well as in its progeny midline glial cells along the dorsal midline of the CNS, but not in the neuronal cell bodies whose axons form the major commissural and longitudinal axon tracts in the CNS. Light and immunoelectron microscopy indicate that SLIT is exported from the midline glial cells and is associated with the axons that traverse them. If, as is suggested by this data, the SLIT gene product is not produced in the neurons of the axons on which it resides, it is expected that it is secreted from the midline cells and "picked up" by passing axons. This in turn raises the possibility that the axons that carry SLIT on their surface may be expressing specific receptors capable of interacting with SLIT in a direct or indirect manner. An analysis of SLIT expressing in Drosophila cell culture demonstrates that SLIT can in fact be localized to the surface of individual cells. Additional biochemical support for the extracellular, secreted nature of the protein was provided by demonstrating that tissue culture cells producing SLIT are secreting the protein into the media. Moreover, consistent with the hypothesis that SLIT functions as an ECM molecule, it was found that the protein to be accumulated in the matrix materials deposited by these cells.

Morphogenetic Regulation of the Neuroepithelium

A model for SLIT function wherein it regulates the morphological differentiation of a cell by attaching to both the ECM and cell surface receptors is consistent with its predicted structure, its expression pattern and phenotype. Like the other ECM glycoproteins, SLIT is composed of repetitive structural motifs and lacks the hydrophilic regions characteristic of membrane-spanning cell-surface adhesion molecules. ECM glycoproteins play a diverse role in development, acting as signals for cell differentiation,

growth and migration. Furthermore, the SLIT-homologous proteoglycan decorin is involved in the control of cell proliferation and has the ability to convert transformed cells to morphological regularity (Yamaguchi, Y., and Ruoslahti, E. (1988). Expression of human proteoglycan in Chinese hamster ovary cells inhibits cell proliferation. Nature 336, 244-246).

SLIT's involvement in the development and differentiation of the midline neuroepithelium and the subsequent formation of commissural axon pathways is demonstrated herein. In a SLIT mutant background the midline cells do not undergo proper differentiation or morphological movements; instead of filling the midline of each neuromere as they do in the wild type embryo, they appear at the base of the nerve cord and are fewer in number. This is followed by the complete collapse of the axonal scaffold. The in vivo effects of reductions in SLIT expression further indicate that the morphogenesis of the midline cells and the subsequent axonal pathway formation are dependent on the concentration of slit protein. Using P-element induced SLIT alleles, applicants were able to demonstrate that a reduction in SLIT expression is coincident with the lack of development of an individual segment's midline cells, and specifically, with the development of the midline glial cells. It was further demonstrated that the variability in the extent of collapse of the midline of the nerve cord is mirrored by the extent of collapse of the commissural and longitudinal axon pathways.

It is noted with interest that the extent of disruption in the ventral nerve cord in slit alleles corresponds to the range of phenotypes exhibited by mutations of the Drosophila EGF-receptor homolog (DER). Given the homology between SLIT and EGF-receptor ligands, the co-localization of the DER and SLIT proteins to the midline glial cells and the muscle

attachment sites (Zak. N. B., Wides, R. J., Schejter, E.D., Raz, E., and Shilo, B. (1990). Localization of the DER/flb protein in embryos: implications on the fait little bal lethal phenotype. Development, 109, 865-874) raises the possibility that SLIT functions as a DER ligand. This speculation is particularly attractive since the activation of a receptor tyrosine kinase by the SLIT protein would offer a mechanistic explanation for SLIT's influence on either the development or maintenance of the midline cells and provide for a direct molecular link between the ECM and genes involved in cellular proliferation and differentiation (Yarden, Y., and Ullrich A. (1988) Growth factor receptor tyrosine kinases. Ann. Rev. Biochem. 57, 443-78).

Implications of SLIT Expression

The three major regions of SLIT expression are the (1) midline neuroepithelium of the central nervous system, (2) the attachment sites of muscle to epidermis, and (3) the cardioblasts of the dorsal tube. The expression of SLIT in the cardioblasts as they meet and form the lumen of the dorsal tube may be of general interest given that, in vertebrate tissue culture, the extracellular matrix has been shown to be involved in endothelial cell alignment and the induction of capillary tube formation. (Ingber, D. E., and Folkman, J. (1989). How Does the Extracellular Matrix Control Capillary Morphogenesis? Cell58, 803-805). This process is one of the best characterized morphogenetic processes in vitro and has allowed for an analysis of the molecular mechanisms by which ECM molecules, specifically collagen, laminin, and fibronectin are able to control capillary morphogenesis (Grant, D. S., Tashiro, K., Segui-Real, B., Yamada, Y., Martin, G. R., Kleinman, H. K. (1989). Two different laminin domains mediate the differentiation of human endothelial cells into capillary-like structures in vitro. Cell 58, 933-943).

In Drosophila, the larval heart or dorsal vessel is derived from two longitudinal rows of mesodermal cells termed cardioblasts. When these cells meet following dorsal closure along the midline, only their dorsomedial and ventromedial surfaces contact, the space between forming the lumen of the dorsal vessel (Poulson, E. F. (1950). Histogenesis, Organogenesis, and Differentiation in the Embryo of Drosophila Melanogaster Meigen. In Biology of Drosophila, M. Demerec, ed. (New York: Wiley), 168-274): Hartenstein, V,., and Campos-Ortega, J. A. (1985). The embryonic development of Drosophila melanogaster. Springer-Verlag; Berlin). SLIT is expressed in the developing cardioblasts during the time they come together. Confocal microscopic imaging clearly shows the SLIT protein to be concentrated at the point of contact between the cardioblasts as they come together and form the lumen of the larval heart. Given SLIT's unique structural characteristics, its homology to ECM binding proteins, and the role of these ECM proteins in vessel formation, an analysis of SLIT's role in developing cardioblasts and its possible interactions with other proteins expressed in these tissues during larval heart formation would serve as a useful in vivo model for the study of the angiogenic process.

Confocal microscopy shows the SLIT protein to be tightly localized to the points of muscle attachment to the epidermis. This localization is consistent with SLIT functioning as an ECM molecule, and suggests its involvement in adhesive events. The muscle attachment sites are known sites of ECM deposition (Newman, S. M. Jr., and Wright, R. F. (1981). A histological and Ultrastructural Analysis of Development Defects Produced by the Mutation, lethal (1) myospheroid, in Drosophila melanogaster. Dev. Bio. 86, 393-402), and the position-specific integrins have been shown to be localized here (Leptin, M., Bogaert, T., Lehmann, R., and Wilcox, M (1989). The Function of PS Integrins during Drosophila

Embryogenesis. Cell 56, 401-408). Hence, a role for SLIT in adhesive-mediated events such as muscle attachment and axonal outgrowth is supported both by its structure and its expression pattern. The potential for two variants of the SLIT protein raises the possibility that these roles are mediated by functionally distinct forms of the protein.

Tissue culture studies have demonstrated that growth cones adhere to and extend neurites onto ECM molecules such as laminin and fibronectin Sanes et al., 1989 supra) and that the direction and rate of axonal growth are dependent on these axon matrix interactions (Rutishauser, U., and Jessell, T. M. (1988). Cell adhesion molecules in vertebrate neural development. Pysiol. Rev. 68, 819-857). Given SLIT's homology to the laminin binding protein biglycan, it is noted with interest that laminin is expressed on glial surfaces and along the pathways axons follow in the establishment of the commissural and longitudinal axonal tracts in Drosophila (Montell, D. J. and Goodman, C. S. (1989). Drosophila laminin: sequence of ß2 subunit and expression of all three subunits during embryogenesis. J. Cell. Bio. 109, 2441-2453). The possibility that SLIT binds to matrix materials suggests that its presence on growing axons could influence their interactions with ECM proteins. The ability of axons to fasciculate on one another in all SLIT mutants indicates that SLIT is not necessary for axon-axon fasciculation. However, the combination of Flank-LRR-Flank tandem EGF and single EGF motifs in a protein with SLIT's unique embryonic distribution could allow for the formation of a "molecular-bridge" between axonally associated receptors and ECM molecules. Prompted by the information on SLIT's structure, its expression in glial cells and its presence on axons which extend along these cells, a mechanism whereby glial cells can influence an axon's future behavior is as follows:

(1) Glial cells secrete multi-functional molecules (TAGONS) into the endoneurial basal lamina. These TAGONS have the ability to attach to specific axonal receptors as well as to specific ECM components.

(2) Passing axons carrying receptors for these proteins pick them up from the glial cell surroundings.

(3) Depending on the proteins associated with them, axons are able to respond to cues and interact with molecules in the ECM.

SLIT is one of the TAGONS.

Therapeutics

The SLIT protein is a unique extracellular matrix protein with applications in nerve regeneration, angiogenesis, and control of neoplasms. SLIT is involved in the development of axon pathways.

The SLIT protein is involved in the development and maintenance of the central nervous system, including the process of glial cell differentiation and neuronal outgrowth. The SLIT protein also plays an inductive role in vessel formation.

The SLIT protein facilitates interactions between cell surface receptors and extracellular matrix molecules, hence providing for a novel molecular link between a cell's environment and genes (including known oncogenes) involved in cellular proliferation and differentiation.

The SLIT protein is involved in the development of cell specificity and the process of neuronal outgrowth.

The SLIT protein molecule can be a therapeutic especially for the repair of damaged neuronal tissue, either alone or in combination with neuronal growth factors (NGF) or other extracellular molecules, and it is useful in nerve repair and tissue regeneration.

The SLIT protein defines a new and novel set of molecules (TAGONS) which play a key role in axon outgrowth and pathfinding. The SLIT protein is thus involved in neurogenesis, axonogenesis, cell differentiation, organ formation and vessel formation and also in muscle attachment.

The SLIT protein can be utilized as a nerve regenerative in neurodegenerative diseases, e.g., it can be utilized as a therapeutic for the following conditions: Alzheimer's disease, spinal cord injuries, brain injuries, crushed optic nerve, nerve damage, amyotrophic lateral sclerosis (ALS), crushed nerves, diabetes-caused nerve damage, facial nerve damage resulting in facial paralysis, Parkinson's disease, strokes, epilepsy, multiple sclerosis, paraplegia and retinal degeneration.

The SLIT proteins of the invention can be formulated into pharmaceutically acceptable preparations with parenterally acceptable vehicles and excipients in accordance with procedures known in the art.

The pharmaceutical preparations of this invention, suitable for parenteral administration, may conveniently comprise a sterile lyophilized preparation of the protein which may be reconstituted by addition of sterile solution to produce solutions, preferably isotonic with the blood of the recipient. The preparation may be presented in unit or multi-dose containers, e.g., in sealed ampoules or vials.

The pharmaceutical preparation may in some instances by orally administered in the form of pills, tablets or capsules.

In use, purified SLIT protein is administered to a mammal, e.g., a human, for treatment in a manner appropriate to the indication. Administration may be by injection, continuous infusion, sustained release from implants (such implants may take the form of a biodegradable plastic or resin having the therapeutic imbedded therein), or other suitable technique. Where the SLIT protein is administered as an aid to wound healing, it will typically be applied topically to the site of the injury, for example, in conjunction with a wound dressing. Therapeutically-effective dosage levels are determined by initiating treatment at higher dosage levels and reducing the amounts of the SLIT protein administered until the condition sought to be alleviated, e.g., wound healing including, but not limited to, neuronal trauma, is no longer achieved. Generally, therapeutic dosages will range from about 0.1 to 1000 ng per g body weight, preferably 1-100 ng/kg. Dosage will vary based on several factors including the weight of the patient and the severity of the condition or ailment. Typically, the SLIT protein will be administered in the form of a composition comprising purified protein in conjunction with physiologically acceptable carriers, excipients or diluents. Neutral buffered saline or saline mixed with conspecific serum albumin are exemplary appropriate diluents.

The present invention also envisages methods for the treatment of animals in need thereof, such animals preferably being mammals, and most preferably being human beings. The treatment will tend to comprise administration of non-toxic formulations described above in the appropriate manner and in suitable doses. SLIT is involved in the development of axon pathways. Alone, or possibly in combination with neuronal growth factors, SLIT is expected to find use in nerve repair and tissue regeneration. The involvement of the SLIT protein

in organ/vessel formation will lead to either direct or indirect therapeutic applications in the control of neoplasms.

Diagnostics

The SLIT proteins according to the present invention and antibodies raised thereto can be employed in immunoassays. Such antibodies can be polyclonal antibodies or monoclonal antibodies.

The detection of SLIT in the bloodstream of a patient is important because such is an indication of an abnormal condition, since SLIT does not appear in the blood of a normal mammal. The presence of SLIT in one's blood may be, for example, an indication of a cancerous condition.

A monoclonal antibody can be prepared according to known methods, for example by the procedures of immunization, cell fusion, screening, and cloning, using the procedures of G. Kohler and C. Milstein (1975), Nature (Lond.), 256, 495.

In selection of the animal to be immunized for production of a monoclonal antibody, the animal species and the immune response to the antigen are important. Generally speaking, stable antibody-producing hybridomas will be frequently formed with good efficiency when the spleen cells to be used and myeloma are of the same animal species. Particularly preferred is the use of BALB/c mice. Preferred myeloma cell species include P3·X63·Ag8(X63), P3·NS-1/1·Ag4·1(NS-1), SP2/O·Ag14(SP-2) and FO.

The antibody, protein or sample in the immunuassays of the invention may be immobilized to a support.

Known immobilization techniques and materials can be employed. Examples of immobilization methods include the

physical adsorption method, the ion bonding method, the covalent bonding method, the support crosslinking method, the support-less crosslinking method, and the inclusion method.

The support may be one generally used, and the choice is not particularly limited. Selection of the support depends on the properties of the material to be immobilized, but it is also necessary to consider the size of particules, the surface area in the three-dimensional network structure, the ratio of hydrophilic sites to hydrophobic sites, chemical composition, strength to pressure, etc. of the support. Typical examples of the support include polysaccharide derivatives such as cellulose, dextran, or agarose; synthetic polymers such as polyacrylamide gel, or polystyrene resin; and inorganic materials such as porous glass, or metal oxide.

With the physical adsorption method, where the material is immobilized by physical adsorption onto a water-insoluble support, examples of particularly preferred supports include inorganic substances such as activated charcoal, porous glass, acidic white clay, bleached clay, kaolinite, alumina, silica gel, bentonite, hydroxyapatite, calcium phosphate, metal oxide, or ceramic; a natural polymer such as starch or gluten; or a porous synthetic resin. Adsorption hydrophobically onto a support having hydrophobic groups such as butyl- or hexyl-"SEPHADEX" is also possible.

With the ion bonding method, where the material is immobilized by binding ionically to a water-insoluble support having ion exchange groups, particularly preferred examples of the support include polysaccharides having ion exchange groups such as DEAE-"SEPHADEX" or synthetic polymer derivatives such as ion exchange resins.

With the covalent bonding method, where the material is immobilized by covalent bonding to a water-insoluble support, examples of particularly preferred supports include those having amino, carboxyl, sulfhydryl, hydroxy, imidazole or phenol groups which are functional groups reactive for instance with diazonium salts, acid azides, isocyanates, or active type alkyl halides.

With the support crosslinking method, where the material is immobilized to the support by covalent binding with the use of a crosslinking reagent such as glutaraldehyde, examples of particularly preferred supports include water-insoluble supports having amino groups, such as AE-cellulose, DEAE-cellulose, partially deacylated chitin, or aminoalkylated porous glass.

With the support-less crosslinking method, where immobilization is effected by crosslinking materials with a reagent having two or more functional groups, no support is particularly required. Examples of preferred crosslinking reagents include glutaraldehyde (forming a Shiff's base), an isocyanic acid derivative (forming a peptide), N,N'-ethylenebismaleimide, bisdiazobenzidine (for diazo coupling), or N,N'-polymethylenebisiodoacetamide (alkylating agent). The material which participates in the crosslinking reaction needs a suitable functional group at the N-end, such as an amino group, phenol group, sulfhydryl group or imidazole group.

With the inclusion method, the method may be classified into the lattice type in which materials to be immobilized are incorporated into fine lattices of polymeric gels, and the microcapsule type in which the antibodies or antigens are coated with semipermeable polymeric films. Examples of preferred supports in the case of the lattice type include polymeric compounds, for example, synthetic polymeric

substances such as polyacrylamide gel, polyvinyl alcohol, or photocurable resin; and natural polymeric substances such as starch, konjak powder, gelatin, alginic acid, or carrageenan. In the case of the microcapsule type, various techniques are possible. When the interfacial polymerization method is used, namely the method in which the antibody is coated by utilizing the principle of polymerizing a hydrophilic monomer and a hydrophobic monomer at the interface therebetween, a nylon film based on hexamethylenediamine and sebacoyl chloride can be employed. When the drying-in-liquid method is used, namely the method in which an antibody solution is dispersed in a polymeric compound solution dissolved in an organic solvent to form an emulsion and then transferred into an aqueous solution followed by drying, thereby coating the antibody, examples of preferred supports include polymeric substances such as ethyl cellulose or polystyrene. When the phase separation method is used, namely the method in which a polymeric compound is dissolved in an organic solvent immiscible with water, an antibody is dispersed in the solution to prepare an emulsion, then a non-solvent which causes phase separation is gradually added under stirring, whereby a concentrated solution of the polymeric compound encloses the antibody droplets therearound, and subsequently the polymeric compound is precipitated to form a film which covers the antibody, is used, the above-mentioned polymeric compounds can be employed.

Labels for use in the present invention include substances which have a detectable physical, chemical or electrical property. When a detectable labeling substance is introduced, it can be linked directly such as by covalent bonds or can be linked indirectly such as by incorporation of the ultimately detectable substance in a microcapsule or liposome.

Labelling materials have been well-developed in the field of immunoassays and in general almost any label useful in such methods can be applied to the present invention. Particularly useful are enzymatically active groups, such as enzymes (see Clin. Chem., (1976) 22:1232, U.S. Reissue Pat. No. 31,006, and UK Pat. 2,019,408), enzyme substrates (see U.S. Pat. No. 4,492,751), coenzymes (see U.S. Pat. Nos. 4,230,797 and 4,238,565), and enzyme inhibitors (see U.S. Pat. No. 4,134,792); fluorescers (see Clin. Chem., (1979) 25:353); chromophores; luminescers such as chemiluminescers and bioluminescers (see U.S. Pat. 4,380,580); specifically bindable ligands such as biotin (see European Pat. Spec. 63,879) or a hapten (see PCT Publ. 83-2286); and radioisotopes such as ^{3}H, ^{35}S, ^{32}P, ^{125}I, and ^{14}C. Such labels are detected on the basis of their own physical properties (e.g., fluorescers, chromophores and radioisotopes) or their reactive or binding properties (e,g., ligands, enzymes, substrates, coenzymes and inhibitors). For example, a cofactor-labeled species can be detected by adding the enzyme (or enzyme where a cycling system is used) for which the label is a cofactor and a substrate or substrates for the enzyme. Such detectable molecule can be some molecule with a measurable physical property (e.g., fluorescence or absorbance) or a participant in an enzyme reaction (e.g., see above list). For example, one can use an enzyme which acts upon a substrate to generate a product with a measurable physical property.

Any convenient immunoassay technique can be employed in the present invention including, for example, enzyme-linked immunoassay, radioimmunoassay (RIA), immunofluorescence and the use of dyes.

In enzyme linked immunoassays, an enzyme is conjugated to an antibody or antigen and the enzyme activity is measured as

a quantitative label. A particularly preferred enzyme linked immunoassay is enzyme-linked immunosorbent assay (ELISA).

The enzyme may be any of the enzymes generally used in enzyme immunoassay, including maleate dehydrogenase, glucose-6-phosphoric acid dehydrogenase, glucose oxidase, peroxidase, acetylcholine esterase, alkali phosphatase, glucoamylase, lysozyme, ß-D-galactosidase, etc., preferably peroxidase, alkali phosphatase or ß-D-galactosidase or horseradish peroxidase.

Immunofluorescence utilizes fluorescent dyes such as fluorescein isothiocyanate or rhodamine.

The detection of nucleic acids involves hybridization conditions and techniques that are known in the art. The principle for the hybridization test is as follows:

Two DNAs are heated to denature them completely, with separation of strands. When they are mixed and slowly cooled, complementary DNAs of each species will find each other and reannel to form normal duplexes. But if the two DNAs have significant sequence homology, they will tend to form partial duplexes or hybrids with each other. The greater the sequence homology between two DNAs, the greater the number of hybrids formed. Hybrid formation can be measured by different procedures, e.g., chromatography or density-gradient centrifugation. Usually one of the DNAs is labeled with a radioactive isotope to simplify the measurements.

The SLIT nucleic acid molecule according to the present invention can be used as a gene probe, i.e., a nucleic acid molecule that can be used to detect, by complementary base-pairing, another nucleic acid molecule that has a complementary or homologous sequence. The probe is invariably

labeled, e.g., Nick translation, Biotin, to allow autoradiographic or enzymatic detection of the hybridization reaction.

The Southern transfer method can be utilized in the present invention. The Southern transfer procedure (developed by Edwin Southern and sometimes called blotting), a method for performing hybridization to particular DNA segments, avoids the necessity of purifying the DNA fragments with restriction endonucleases.

At present the best way to separate DNA fragments from one another is by electrophoresis through agarose gels. A specific fragment can be isolated by cutting out of a gel a portion that contains the fragment of interest. A variety of procedures, most of which are cumbersome and tedious, are available for recovering the DNA molecule from the gel. If hybridization is to be performed, the fragment must be bound to a nitrocellulose filter. In the Southern transfer technique a collection of fragments is handled in such a way that all fragments are transferred from a gel to a sheet of nitrocellulose in a single step, significantly simplifying the entire process.

The Southern transfer technique is carried out as follows DNA is enzymatically fragmented and then electrophoresed through an agarose gel. Following electrophoresis the gel is soaked in a denaturing solution (usually NaOH), so that all DNA in the gel is converted to single-stranded DNA, which is needed for hybridization. A large sheet of nitrocelulose paper is placed on top of several sheets of ordinary filter paper; the gel, which is typically in the form of a broad flat slab, is then placed on the nitrocellulose filter and covered with a glass plate to prevent drying. A weight is then placed on the top of the stack and the liquid is squeezed out of the

gel. The liquid passes downward through the nitrocellulose filter. Denatured DNA binds tightly to nitrocellulose; the remaining liquid passes through the nitrocelulose and is absorbed by the filter paper. DNA molecules do not diffuse very much, so that if the gel and the nitrocellulose are in firm contact, the positions of the DNA molecules on the filter are identical to their positions in the gel. The nitrocellulose filter is then dried in vacuum, which insures that the DNA remains on the filter during the hybridization step. The dried filter is then moistened with a very small volume of a solution of ^{32}P-labeled RNA, placed in a tight-fitting plastic bag to prevent drying, and held at a temperature suitable for renaturation (usually for 16-24 hours). The filter is then removed, washed to remove unbound radioactive molecules, dried, and auto radiographed with x-ray film. The blackened positions of the film indicate the locations of the DNA molecules whose DNA base sequences are complementary to the sequences of the added radioactive molecules.

For the most part, the probe will be labeled with an atom or inorganic radical, most commonly using radionuclides, but also perhaps heavy metals.

Conveniently, a radioactive label may be employed. Radioactive labels include ^{32}P, ^{3}H, ^{14}C, or the like. Any radioactive label may be employed which provides for an adequate signal and has sufficient half-like. Other labels include ligands, which can serve as a specific binding member to a labeled antibody, fluorescers, chemiluminescers, enzymes, antibodies which can serve as a specific binding pair member for a labeled ligand, and the like. A wide variety of labels have been employed in immunoassays, as discussed hereinabove, which can readily be employed in the present hybridization assay. The choice of the label will be governed by the effect

of the label on the rate of hybridization and binding of the probe to the genetic nucleic acid, e.g., DNA. It will be necessary that the lable provide sufficient sensitivity to detect the amount of DNA available for hybridization. Other considerations will be ease of synthesis of the probe, readily available instrumentation, ability to automate, convenience, and the like.

The manner in which the label is bound to the probe will vary depending upon the nature of the label. For a radioactive label, a wide variety of techniques can be employed. Commonly employed is Nick translation with an α-^{32}P-dNTP or terminal phosphate hydrolysis with alkaline phosphatase followed by labeling with radioactive ^{32}P employing γ-^{32}P-NTP and T4 polynucleotide kinase. Alternatively, nucleotides can be synthesized where one or more of the elements present are replaced with a radioactive isotope, e.g., hydrogen with tritium. If desired, complementary labeled strands can be used as probes to enhance the concentration of hybridized label.

Where other radionuclide labels are involved, various linking groups can be employed. A terminal hydroxyl can be esterified, with inorganic acids, e.g., ^{32}P phosphate, or ^{14}C organic acids, or else esterified to provide linking groups to the label. Alternatively, intermediate bases may be substituted with activatable linking groups which can then be linked to a label.

Ligands and antiligands may be varied widely. Where a ligand has a natural receptor, namely ligands such as biotin, thyroxine, and cortisol, these ligands can be used in conjunction with labeled naturally occurring receptors. Alternatively, any compound can be used, either haptenic or antigenic, in combination with an antibody.

Enzymes of interest as labels will primarily be hydrolases, particularly esterases and glycosidases, or oxidoreductases, particularly peroxidases. Fluorescent compounds include fluorescein and its derivatives, rhodamine and its derivatives, dansyl, umbelliferone, etc. Chemiluminescers include luciferin, and 2,3-dihydrophthalazinediones, e.g., luminol.

The probe can employed for hybridizing to a gene affixed to a water insoluble porous support. The single stranded nucleic acid is affixed. Depending upon the source of the nucleic acid, the manner in which the nucleic acid is affixed to the support may vary.

A clinical isolate or specimen can be spotted or spread onto a filter to provide a plurality of individual portions. The filter is an inert porous solid support, e.g., nitrocellulose. The clinical isolate can be blood or another bodily fluid from a patient, e.g., a human patient. Conveniently, a microfilter is employed, which inhibits the passage of the cells through the filter.

The cells are then treated to liberate their DNA. Lysis conditions are devised such that the cells do not migrate and their DNA remains affixed in place on the surface where they were situated. The lysing and DNA denaturing, as well as the subsequent washings, can be achieved by placing the filter containing the cells isolate side up, onto a bibulous support saturated with an appropriate solution for a sufficient time to lyse the cells and denature the DNA. For lysing, chemical lysing will conveniently be employed, usually dilute aqueous alkali, e.g., 0.1 to 1 M NaOH. The alkali will also serve to denature the DNA. Other denaturation agents include, elevated temperatures, organic reagents, e.g., alcohols, amides,

amines, ureas, phenols and sulfoxides or certain inorganic ions, e.g., thiocyanate and perchlorate.

After denaturation, the filter is washed in an aqueous buffered solution, generally at a pH of about 6 to 8, usually 7. Of the many different buffers that may be used, tris is an example. One or more washings may be involved, conveniently using the same procedure as employed for the lysing and denaturation.

After the lysing, denaturing and washes have been accomplished, the DNA spotted filter is dried at an elevated temperature, generally from about 50° to 70°C. The DNA is now fixed in position and can be assayed with the probe when convenient. This fixing of the DNA for later processing has great value for the use of this technique in field studies, remote from laboratory facilities.

Hybridization may now be accomplished. The filter is incubated at a mildly elevated temperature for a sufficient time with the hybridization solution without the probe to thoroughly wet the filter. Various hybridization solution may be employed, comprising from about 20 to 60 volume, preferably 30, percent of an inert polar organic solvent. A common hybridization solution employs about 50% formamide, about 0.5 to 1 M sodium chloride, about 0.05 to 0.1 M sodium citrate, about 0.05 to 0.2% sodium dodecylsulfate, and minor amounts of EDTA, ficoll (about 300-500 kdaltons), polyvinylpyrrolidone, (about 250-500 kdaltons) and serum albumin. Also included in the hybridization solution will generally be from about 0.5 to 5 mg/ml of sonicated denatured DNA, e.g., calf thymus of salmon sperm; and optionally from about 0.5 to 2% wt/vol. glycine. Other additives may also be included, such as dextran sulfate of from about 100 to 1,000 kdaltons and in an amount

of from about 8 to 15 weight percent of the hybridization solution.

The particular hybridization technique is not essential to the invention. Other hybridization techniques are described by Gall and Pardue (1969) Proc. Natl. Acad. Sci, 63, 378-383 and John, Burnsteil and Jones, Nature, 223, 582-587, (1969). As improvements are made in hybridization techniques they can readily be applied.

The amount of labeled probe which is present in the hybridization solution will vary widely, depending upon the nature of the label, the amount of the labeled probe which can reasonably bind to the filter, and the stringency of the hybridization. Generally, substantial excesses over the stoichiometric amount of the probe will be employed to enhance the rate of binding of the probe to the fixed DNA.

Various degrees of stringency of hybridization will be employed. The more severe the conditions, the greater the complementarity that is required for hybridization between the probe and the ssDNA (single stranded DNA) for duplex formation. Severity can be controlled by temperature, probe concentration, probe length, ionic strength, time, and the like. Conveniently, the stringency of hybridization is varied by changing the polarity of the reactant solution by manipulating the concentration of formamide in the range of 20% to 50%. Temperatures employed will normally be in the range of about 20° to 80°C, usually 30° to 75°C.

After the filter has been contacted with a hybridization solution at a moderate temperature for an extended period of time, the filter is then introduced into a second solution having analogous concentrations of sodium chloride, sodium citrate and sodium dodecylsulfate as provided in the

hybridization solution. The time for which the filter is maintained in the second solution may vary five minutes to three hours or more. The second solution determines the stringency, dissolving cross duplex and short complementary sequences. After rinsing the filter at room temperature with dilute sodium citrate-sodium chloride solution, the filter may now be assayed for the presence of duplexes in accordance with the nature of the label. Where the label is radioactive, the filter is dried and exposed to X-ray film.

Expression

The nucleotide sequences of the invention are preferably sequences of DNA. Such sequences may be used alone, for example as probes, but it is generally preferred that they form part of an expression system. Thus, it is preferred that the DNA sequence form part of a vector useful in an expression system.

The general nature of vectors for use in accordance with the present invention is not crucial to the invention. In general, suitable vectors and expression vectors and constructions therefor will be apparent to those skilled in the art.

Suitable expression vectors may be based on phages or plasmids, both of which are generally host-specific, although these can be engineered for other hosts. Other suitable vectors include cosmids and retroviruses, and any other vehicles, which may or may not be specific for a given system. Again, control sequences, such as recognition, promoter, operator, inducer, terminator and other sequences essential and/or useful in the regulation of expression, will be readily apparent to those skilled in the art, and may be associated with the natural SLIT protein sequence or with the vector

used, or may be derived from any other source as suitable. The vectors may be modified or engineered in any suitable manner.

In general, there are a number of methods which can be used to produce the peptide and nucleotide sequences of the invention. One straightforward method is simply to synthesize the appropriate nucleotide sequence, insert it into a suitable expression plasmid, transform a suitable host, culture the host, and obtain the SLIT protein of the invention by any suitable means, such as sonication and centrifugation.

Alternatively, fragments can be obtained by digestion with the relevant restriction enzymes, and a suitable oligonucleotide ligated to the 5'-end coding for missing amino acids. The resulting cDNA can then be used as above.

Other suitable methods will be apparent to those skilled in the art.

It will be appreciated that the fragment encoding the SLIT protein of the invention may easily be inserted into any suitable vector for any purpose desired. Suitable vectors may be selected as a matter of course by those skilled in the art according to the expression system desired.

By transforming E. coli with the plasmid obtained, selecting the transformant with ampicillin or by other suitable means, and adding tryptophan or other suitable promoter inducer such as indoleacrylic acid, the desired protein may be expressed. The extent of expression may be analyzed by SDS polyacrylamide gel electrophoresis - SDS-PAGE (Nature, (1970), 227, pp.680-685).

It will also be appreciated that, where another vector is used, for example, it will be equally acceptable to employ a different selection marker or markers, or an alternative method of selection, and/or to use any suitable promoter as required or convenient.

After cultivation, the transformant cells are suitably collected, disrupted, for example, sonicated, and spun-down. Disruption may also be by such techniques as enzymic digestion, using, for example, cellulase, or by shaking with an agent such as glass beads, but methods such as sonication are generally preferred, as no additions are necessary. The activity of the supernatant may be assayed and the amount of the SLIT protein measured by SDS-PAGE, for example, allowing the specific activity to be calculated.

Conventional protein purification is suitable to obtain the expression product.

Where not specifically described herein, methods for growing and transforming cultures etc. are usefully illustrated in, for example, Maniatis (Molecular Cloning, A Laboratory Notebook, Maniatis et al. [Ed's], Cold Spring Harbor Labs, NY).

Cultures useful for the production of the SLIT protein of the invention may suitably be cultures of any living cells, and may vary from prokaryotic expression systems up to eukaryotic expression systems. One preferred prokaryotic system is that of E. coli, owing to its ease of manipulation. However, in general terms, it is preferable to express proteins intended for use in the human body in higher systems, especially mammalian cell lines. A currently preferred such system is the Chinese Hamster Ovary (CHO) cell line. Although this system tends not to be as easy to use as the E. coli

system, its advantage lies in the processing of the protein after primary synthesis. E. coli, for example, does not have the equipment to glycosylate mammalian proteins, and it is preferred to glycosylate such proteins where possible, if for no other reason than that the natural proteins are glycosylated. In certain cases, glycosylation may be of no assistance whatever, and may even hinder the protein. In the present instance, glycosylation appears to serve little purpose.

Other expression systems which may be employed include streptomyces, for example, and yeasts, such as Saccharomyces spp., especially S. cerevisiae. With current progress in research, other systems are becoming available and there is no effective limit on which system is used, provided that it is suitable. The same systems may also be used to amplify the genetic material, but it is generally convenient or use E. coli for this purpose where only proliferation of the DNA is required.

Equivalents, Variants and Mutants

In general, it will be appreciated that the activity of any given protein is dependent upon certain conserved regions of the molecule, while other regions have little importance associated with their particular sequence, and may be virtually or completely redundant. Accordingly, the present invention also includes any equivalents, variants and mutants on the sequence which still show substantial activity. Such variants and mutants include, for example, deletions, insertions, repeats and type-substitutions (e.g., substituting one hydrophilic residue for another, but not strongly hydrophilic for strongly hydrophobic as a rule). Small changes will be generally have little effect on activity, unless they are an essential part of the molecule, and may be

a side-product of genetic manipulation, for example, when generating extra restriction sites, if such is desired.

It will be appreciated that the coding sequence may be modified in any manner desired, provided that there is no adverse effect on activity. Spot mutations and other changes may be effected to add or delete restriction sites, for example, to otherwise assist in genetic manipulation/expression, or to enhance or otherwise conveniently modify the SLIT molecule.

As used herein, the term a "adverse effect" means any effect on activity, or as otherwise used, which renders the molecule only as effective as, or less effective as, the naturally occurring SLIT protein.

If desired, the carboxy terminal group or other carboxyl groups of the SLIT protein may be substituted or modified in any manner apparent to those skilled in the art. Such substitutions may include the formation of salts and esters, for example, or any other substitution as appropriate. Modification may include the deletion of one or more C-terminal amino acid residues, partially or entirely, provided that this has no adverse effect on activity. Deletion of the terminal carboxyl group may be useful in preventing undesirable reactions, which purpose may also be served by the use of an appropriate protecting group, for example. Modification may also include replacement of one or more of the residues with any other suitable residue, and such replacement may either be 1:1 or any other suitable ratio.

Modifications but, more especially, substitutions to the C-terminal may either be temporary or permanent, as with modifications and substitutions to the SLIT protein molecule as a whole. Thus, a C-terminal esterified SLIT protein may be

de-esterified in vivo, either at or before reaching the target site. Likewise, the SLIT protein may be specifically modified, particularly by deletion or substitution, so as to be inactive until the target is reached, whereon activation may be internal, by enzymatic cleavage or addition, for example, or external, such as by irradiation to activate a sensitive group.

In general, it will be appreciated that the entire molecule may be substituted or modified within wide limits. Thus, for example, it will be apparent that the SLIT protein of the invention may be heavily glycosylated without adversely affecting activity. The present invention envisages both glycosylated and unglycosylated SLIT protein of the invention as being useful, as well as any state in between.

Many substitutions, additions, and the like may be effected, and the only limitation is that activity not be adversely affected. In general, an adverse effect on activity is only likely if the 3-D (tertiary) structure of the SLIT protein is seriously affected, or if an active site is somehow affected (reducing electronegativity/hydrophilicity, blocking etc.).

If it is desired to glycosylate the SLIT protein molecule selectively, rather than randomly as would be achieved by direct chemical addition, this can be achieved best by a eukaryotic, especially mammalian, system. This may either comprise a eukaryotic expression system, or treatment of the product with a suitable enzyme system in vitro, both of which are known in the art.

Selective substitution on the molecule will not generally be facile. For example, to modify only the C-terminal carboxy, it would most likely be necessary to block any other

groups likely to be modified by the same treatment. Universal modification of a particular type of group may be acceptable, such as esterification, but it is usually acceptable and, moreover, practical to use the unmodified expression product. However, selective modification is particularly achieved by appropriate selection of expression system and/or suitable modification of the coding sequence.

Suitable substitutions, additions and the like may be effected as desired to assist in formulation, for example, or may be a product of any expression system employed.

With reference to the peptide sequences disclosed herein, the term "equivalent" is used in the sense of the preceding description, that is to say, equivalents in the sense of sequences having substitutions at the C- or N-terminals, or anywhere else, including salts and esters, and glycosylated sequences. The term "mutants" is used with reference to deletions, insertions, inversions and replacements of amino acid residues in the sequence which do not adversely affect activity. "Variant" is used in relation to other naturally occurring SLIT proteins which may be discovered from time to time and which share essentially as shown in the sequences herein, but which vary therefrom in a manner to be expected within metoazoan organisms. Within this definition lies allelic variation. The term "precursor" includes such molecules as those having leader sequences or substitutions which may or may not affect activity, but which are no longer present when the SLIT protein is active, whether the effect was negated before or at the target site.

The present invention also provides nucleotide sequences encoding all or part of the SLIT proteins of the invention. As will be apparent from the foregoing, there is little restriction on the sequence, whether it be DNA or RNA. A gene

encoding the SLIT proteins of the invention may easily be reverse-engineered by one skilled in the art from the sequences given herein together with the information provided herein.

It will be appreciated that any one given reverse-engineered sequence will not necessarily hybridize well, or at all, with any given complementary sequence reverse-engineered from the same peptide, owing to the degeneracy of the genetic code. This is a factor common in the calculations of those skilled in the art, and the degeneracy of any given sequence is frequently so broad as to make it extremely difficult to synthesize even a short complementary oligonucleotide sequence to serve as a probe for the naturally occurring oligonucleotide sequence.

The degeneracy of the code is such that, for example, there may be four, or more, possible codons for frequently occurring amino acids. Accordingly, therefore, it can be seen that the number of possible coding sequences for any given peptide can increase exponentially with the number of residues. As such, it will be appreciated that the number of possible coding sequences for the SLIT protein of the invention may have several figures, with little to choose between any of that number. However, it may be desirable to balance the GC ratio according to the expression system concerned, and other factors may need to be taken into account which may affect the choice of coding sequence.

The invention is now described with reference to the following non-limiting examples.

Example 1: **Cloning By Transposon Tagging**

$slit^{F81}$ and $slit^{F119}$ were created by germline transformation with the enhancer trap construct P-lacW (Bier et al., (1989).

Searching for pattern and mutation in the Drosophila genome with a P-lacZ vector. Genes & Dev. 3, 1273-1287) and $slit^{E158}$ was made using P-lArB (Bellen, H. J., O'Kane, C. J., Wilson, C., Grossniklaus, U., Pearson, R. K., and Gehring, W. J. (1989). P-element-mediated enhancer detection: a versatile method to study development in Drosophila. Gen. & Dev. 3, 1288-1300). Other SLIT alleles are as described in Rothberg et al., 1988, supra $slit^{175}$ exhibit some ectopic ß-galactosidase expression, while $slit^{F81}$ and $slit^{F119}$ (likely the result of the same insertion event) have levels of midline expression lower than levels in $slit^{E158}$. Lines 8-7 and 242 function as cell autonomous markers for the midline neuroepithelium and contain the PZ and HZ enhancer trap constructs which use the P-element and ftz promoters, respectively, to drive ß-galactosidase expression. Line 5704 expresses ß-galactosidase from the ftz promoter in the MP2 cells (Hiromi, Y., Kuroiwa, A., and Gehring, W. J. (1985). Control elements of the Drosophila segmentation gene fushi tarazu. Cell 43,603-613). Lines 8-7, 242 and 5704 were made homozygous in $slit^{IG107}$/CyO flies to characterize the development of the midline in $slit^{IG107}$/$slit^{IG107}$ embryos.

Example 2: **Isolation of cDNA and Genomic Clones**

Isolation of the initial slit cDNA clones was described in Rothberg et al., (1988), supra. Both the polymerase chain reaction (PCR) (Saiki R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R. Horn, G. T., Mullis, K. B., and Erlich, H. A. 1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239, 487-491) and standard library screening methods (Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982). Molecular Cloning: A Laboratory Manual (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory) were employed to extend this analysis. A cDNA clone representing the 5' most 2.4kb of sequence (ka2.4) was isolated from the larval library of Poole, S. J. Kauvar,

L. M. Drees, B., and Kornberg, T. (1985). The engrailed locus of Drosophila: structure analysis of an embryonic transcript. Cell 40, 37-43) and PCR was used to isolate a corresponding sequence (be2.4) from a 4-8 hour embryonic library (Brown, N. M. and Kafatos, F. C.(1988), J. Mol. Biol. 203, 425-437).

Two forms of the SLIT message were evident differing by 33 nucleotides, when restriction fragments from the larger class (B52-1 and B52-2) were compared with those from the smaller class (B52-5). Primer pairs covering adjacent segments of the coding region were utilized in the PCR to screen embryonic cDNA libraries (Poole et al., 1985, supra; Brown and Kafatos, 1988, supra) for the presence of multiple cDNA forms. Two classes already represented by B52-1,2 and B52-5 were generated. Genomic and cDNA sequencing indicates the transcripts consists of an approximately 314 bp 5' untranslated leader sequence, followed by either a 4407 or 4440 bp ORF depending on the splice form and a 4 kb untranslated 3' end. EcoRI cDNA fragments representing the entire transcription unit were aligned with genomic sequences by Southern analysis.

Example 3: **Subcloning, Sequencing, Localization of Transposon insertion Sites**

The relevant regions from phage, plasmid and PCR-generated cDNAs were subcloned into Bluescript (Stratagene) or M13mp18/19 vector. Single-stranded templates were sequenced directly or subjected to deletions by T4 polymerase (International Biotechnologies Corp.). Chain termination sequencing (Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74, 5463-5467) used Sequenase v2.0 (United States Biochemical Corp.). dITP was employed where sequence was ambiguous and synthetic

oligonucleotides were used as primers to fill any gaps in the nested deletions. The use of gene-specific and P-element inverted repeat-specific primers to isolate genomic DNA using PCR was previously described in Ballinger, D. G., Benzer, S.,(1990). Targeted gene mutations in Drosophila. Proc. Natl. Acad. Sci. USA 86, 9402-9406. Sequences from the 31 bp inverted P-element repeat (O'Hare, K., and Rubin, G.M. (1983). Structure of P Transposable Elements and Their Sites of Insertion and Excision in the Drosophila melanogaster Genome. Cell 34, 25-35) and from the 5' region of the SLIT transcript were used as primers. Sequencing of PCR products was performed on a Dupont Genesis 2000 sequencing machine after the generation of single-stranded DNA by asymmetric PCR and the removal of excess primers with Sepharose S-200 spin columns. Sequence analysis was accomplished with MacVector (International Biotechnologies Inc.) on a Macintosh II. Database searches and sequence comparisons were conducted using the FASTA package (Pearson, W. R., and Lipman, D. J.. (1988). Improved tools for biological sequence comparison. Proc. Nat. Acad. Sci. USA 85, 2444-2448) with version 23 of the NBRF database.

Example 4: **Whole Mount in situs, Enhancer Trap Detection and Antibody Labeling**

Whole mount in situ hybridizations were conducted using digoxigenin-derivatized DNA probes from cDNA B52-5. Immunocytochemistry was done essentially as described in Rothberg et al, 1988, supra. Anti-ß-galactosidase antibody (Promega Corp.) was used to detect the signal from the enhancer trap constructs and detected with a horseradish peroxidase (HRP)-conjugated anti-mouse antibody (Jackson Immunological Laboratories). Signal from whole mount in situ is cytoplasmic (Tautz, D., and Pfeiffle, C. (1989). A non-radioactive in situ hybridization method for the localization of specific RNAs in Drosophila embryos reveals transnational

control of the segmentation gene hunchback. Chromosoma 98, 81-85) enhancer trap signal is localized to the nucleus (Bellen et al., 1990), and antibody staining shows both cytoplasmic and cell surface staining.

Example 5: **Immunoelectron and Confocal Microscopy**

All preparations were made by dissecting embryos in Schneider medium to expose the nerve cord. Samples were fixed in 2% paraformaldehyde with .025% glutaraldehyde for 15 minutes followed by primary and secondary antibody labeling without detergent. Primary E.M. fixation was performed using 2% glutaraldehyde and 2% paraformaldehyde prior to silver enhancement of signal from the HRP-conjugated secondary (Amersham Corp). The silver enhancement procedure prevents accurate distinctions to be made concerning the relative levels of antigen present among subsets of axons. Samples were treated with 1% OsO_4 and counter-stained with Uranyl-acetate. Sections were prepared on a Reichert ultramicrotome and visualized on a Jeol electron microscope. Confocal images were made using a Biorad MRC 500 system and a Zeiss Axiovert compound microscope.

Example 6: **Immunofluorescence, Immunoprecipitations, and Immunoblots**

Immunofluorescence of Drosophila S2 cell lines, the preparation of lysates from Canton-S embryos and S2 cell lines (Schneider, I. (1972). Cell lines derived from late embryonic stages of Drosophila melanogaster. J. Embryol. exp. Morph 27, 353-365) were performed essentially as described in Fehon, R. G., Kooh, P. J., Rebay, I., Regan C. L., Xu, T., Muskavitch, M. A. T., and Artavanis-Tsakonas, S. (1990). Molecular interactions between the protein products of the neurogenic loci Notch and Deta, two EGF-homologous genes in Drosophila. Cell 61, 523-534. Immunoprecipitation of protein lysates and S2 cell conditioned media were performed with anti-slit

antibodies followed by the precipitation of the immune complex with protein A-sepharose 6MD (Pharmacia) or protein A/G beads (Pierce). Samples were suspended in SDS-PAGE loading buffer, boiled, and separated by SDS-PAGE. Following transfer to nitrocellulose, blots were probed with anti-slit antibodies and detected with HRP-conjugated goat anti-rabbit antibodies. No immunoprecipitatable species from KC cell lysates or conditioned media was detected by immunoblotting matrix proteins deposited by S2 cells grown in plastic culture flasks (T75; Corning) were prepared, after removal of the cells and 3 rinses with 1X PBS, by directly boiling in 300-500 μl of SDS-PAGE loading buffer. 5-10μl were used per lane for immunoblot analysis. Detection of S^{35} labeled slit protein in the media was performed by metabolically labelling (0.1mCi/ml, ICN translabel) S2 cells for 4 hours in M3 media (minus methionine and cysteine), followed by immunoprecipitating the conditioned media with anti-SLIT antibody and protein A-sepharose 6MD. Precipitates were washed overnight in PBS with 1% bovine serum albumin and 0.1% TRITON followed by separation with SDS-PAGE and autoradiography.

Example 7: **Purification of Pure and Active SLIT protein**

Conditioned media from tissue culture cells expressing the natural form of the SLIT protein or detergent extracts of protein lysates expressing SLIT are passed through an antibody column consisting of anti-SLIT IgG monoclonal antibody coupled to Sepharose CL beads (10 mg Mab/ml swollen beads). The column is then washed with 10 bed volumes of PBS and 0.1% TRITON. The protein is then eluted using a 50 mM diethylamine-HCl pH 11.5 and 0.5% deoxychloate buffer and neutralized with glycine. The eluted fractions are monitoring by antigenic activity and shown to be in pure form by SDS-PAGE. The biological activity of the protein is monitored by an axonal outgrowth assay. The same procedure is used to isolate and assay recombinant forms of the SLIT protein

consisting of the various sequence elements defined in this application. Stable Drosophila cell lines over expressing the SLIT protein were constructed by cloning the coding portions of the SLIT gene into the metallothionein promoter vector pRmHa-3 (Bunch et al, 1988 et al., Characterization and use of the Drosophila metallothionein promoter in cultured Drosophila melanogaster cells. Nucl. Acids Res. 16, 1043-1061) and transfecting into the S2 cell lines (Schneider, 1972).

It will be appreciated that the instant specification is set forth by way of illustration and not limitation, and that various modifications and changes may be made without departing from the spirit and scope of the present invention.

SEQUENCE LISTING

(1) GENERAL INFORMATION:
(i) Applicant: Rothberg, Jonathan Marc and Artavanis-Tsakonas, Spyridon
(ii) TITLE OF INVENTION: Purified SLIT protein and Sequence Elements Thereof
(iii) NUMBER OF SEQUENCES: 9
(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSE: Yale University
Office of Cooperative Research
(B) STREET: 246 Church Street
Suite 401
(C) CITY: New Haven
(D) STATE: Connecuticut
(E) COUNTRY: USA
(F) ZIP: 06510
(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette, 3.50 inch. 800 Kb storage
(B) COMPUTER: Apple Macintosh
(C) OPERATING SYSTEM: Macintosh 6.0.5
(D) SOFTWARE: Microsoft Word 4.0
(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: 07/624,135
(B) FILLING DATE: 7-DEC-1990
(C) CLASSIFICATION:
(vii) PRIOR APPLICATION DATA: not applicable
(viii) ATTORNEY INFORMATION:
(A) NAME: Barth, Richard J.
(B) REGISTRATION NUMBER: 28,180
(C) REFERENCE/DOCKET NUMBER: 900964/RSB
(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (212) 972-1400
(B) TELEFAX: (212) 370-1622
(C) TELEX: 236268

(2) INFORMATION FOR SEQ ID NO: 1:
(i) SEQUENCE CHARACTERESITICS:
(A) LENGTH: 8378
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA
(iii) HYPOTHETICAL: no
(iv) ANTI-SENCE: no
(vi) ORIGINAL SOURCE:
(A) ORGANISM: Drosophila Melanogaster
(D) DEVELOPMENTAL STAGE: embryonic and larval, germ-line derived.
(vii) IMMEDIATE SOURCE:
(A) LIBRARY: cDNA and Genomic
(B) CLONE: be2.4, ka2.4, B52-2, B52-5, smart2-19
(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: 2R
(B) MAP POSITION: 52D
(C) UNITS: chromosome band
(ix) FEATURE:
(A) NAME/KEY: 5' leader sequence
(B) LOCATION: 1 to 314
(C) IDENTIFICATION METHOD: experimental

(A) NAME/KEY: Translated region
(B) LOCATION: 315 to 4754
(C) IDENTIFICATION METHOD: experimental
(D) OTHER INFORMATION: codes for slit protein

(A) NAME/KEY: 3' untranslated region
(B) LOCATION: 4755 to 8378
(C) IDENTIFICATION METHOD: experimental

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

TCAGTTTGGT CAAGAAGCGC GTTCGCGACG GCTGCAAAAG AGCGTACCGC CGTAGGAAAA 60

CCCCGAGAGA AAAGTGCGCC GTGGAGCCGG GCGGACATTC ACCGAACCCA AAACGCCTCG 120

AACTCGATAT CGAATCGAAA GGATTAATCC AGTGAAATCA GTGAAGTGAA AGTGCCTGCG 180

AACGCATCAT CAATCCTTTA TCCTTTCTCC CTCAAATATT TACCCAGTGG TGATTGCTGT 240

TGACAAAGTG GATTGGCATA TACGGGGGCC ACTTTCAATT AGGCACGTTG CCGCTGCTTC 300

ATAAATGTGC CACA 314

ATG GCC GCG CCG TCC AGG ACG ACG TTG ATG CCA CCA CCA TTC CGG 359
Met Ala Ala Pro Ser Arg Thr Thr Leu Met Pro Pro Pro Phe Arg
5 10 15

CTC CAG CTG CGG CTA CTG ATA CTA CCC ATC CTG CTA CTC CTG CGC 404
Leu Gln Leu Arg Leu Leu Ile Leu Pro Ile Leu Leu Leu Leu Arg
20 25 30

```
CAT GAT GCG GTC CAC GCG GAA CCG TAT TCC GGC GGA TTC GGC AGC  449
His Asp Ala Val His Ala Glu Pro Tyr Ser Gly Gly Phe Gly Ser
                35                  40                  45

TCA GCT GTA TCC AGC GGT GGA CTG GGG TCA GTG GGC ATT CAC ATA  494
Ser Ala Val Ser Ser Gly Gly Leu Gly Ser Val Gly Ile His Ile
                50                  55                  60

CCC GGC GGC GGA GTG GGC GTC ATC ACG GAG GCC CGC TGC CCG AGG  539
Pro Gly Gly Gly Val Gly Val Ile Thr Glu Ala Arg Cys Pro Arg
                65                  70                  75

GTC TGC TCC TGC ACC GGA TTA AAT GTG GAT TGC TCG CAT CGA GGA  584
Val Cys Ser Cys Thr Gly Leu Asn Val Asp Cys Ser His Arg Gly
                80                  85                  90

CTC ACC TCC GTT CCC AGG AAA ATC TCA GCG GAC GTG GAG CGA CTC  629
Leu Thr Ser Val Pro Arg Lys Ile Ser Ala Asp Val Glu Arg Leu
                95                 100                 105

GAG CTG CAG GGA AAC AAT TTG ACC GTG ATA TAC GAG ACG GAT TTC  674
Glu Leu Gln Gly Asn Asn Leu Thr Val Ile Tyr Glu Thr Asp Phe
               110                 115                 120

CAG CGG CTG ACC AAG CTG CGA ATG CTC CAA CTA ACT GAC AAT CAG  719
Gln Arg Leu Thr Lys Leu Arg Met Leu Gln Leu Thr Asp Asn Gln
               125                 130                 135

ATC CAC ACG ATC GAG AGG AAC TCC TTC CAA GAT TTG GTC TCA CTC  764
Ile His Thr Ile Glu Arg Asn Ser Phe Gln Asp Leu Val Ser Leu
               140                 145                 150

GAG CGA CTG GAC ATC TCC AAC AAT GTC ATC ACG ACC GTG GGT AGA  809
Glu Arg Leu Asp Ile Ser Asn Asn Val Ile Thr Thr Val Gly Arg
               155                 160                 165

CGC GTC TTC AAG GGA GCC CAA TCG TTG CGG AGT CTT CAG CTG GAC  854
Arg Val Phe Lys Gly Ala Gln Ser Leu Arg Ser Leu Gln Leu Asp
               170                 175                 180

AAT AAC CAA ATC ACC TGC CTG GAT GAG CAC GCC TTT AAG GGA TTG  899
Asn Asn Gln Ile Thr Cys Leu Asp Glu His Ala Phe Lys Gly Leu
               185                 190                 195

GTG GAG CTG GAG ATA CTC ACG CTG AAC AAC AAC AAC CTG ACT TCC  944
Val Glu Leu Glu Ile Leu Thr Leu Asn Asn Asn Asn Leu Thr Ser
               200                 205                 210

CTG CCG CAC AAC ATC TTC GGC GGA CTG GGA CGT TTG CGG GCA CTC  989
Leu Pro His Asn Ile Phe Gly Gly Leu Gly Arg Leu Arg Ala Leu
               215                 220                 225

CGG CTG TCG GAC AAT CCG TTC GCC TGC GAC TGC CAT CTG TCC TGG 1034
Arg Leu Ser Asp Asn Pro Phe Ala Cys Asp Cys His Leu Ser Trp
               230                 235                 240
```

```
CTG TCG CGA TTC CTT CGC AGT GCC ACC CGC CTG GCG CCC TAC ACC  1079
Leu Ser Arg Phe Leu Arg Ser Ala Thr Arg Leu Ala Pro Tyr Thr
                245                 250                 255

CGC TGC CAG TCG CCA TCG CAG CTG AAG GGC CAA AAC GTG GCG GAC  1124
Arg Cys Gln Ser Pro Ser Gln Leu Lys Gly Gln Asn Val Ala Asp
                260                 265                 270

CTG CAC GAC CAG GAG TTC AAA TGC TCG GGT CTG ACG GAG CAC GCA  1169
Leu His Asp Gln Glu Phe Lys Cys Ser Gly Leu Thr Glu His Ala
                275                 280                 285

CCG ATG GAA TGC GGG GCG GAG AAC AGC TGT CCG CAC CCA TGT CGC  1214
Pro Met Glu Cys Gly Ala Glu Asn Ser Cys Pro His Pro Cys Arg
                290                 295                 300

TGT GCG GAC GGG ATC GTC GAT TGC CGT GAG AAG AGT CTG ACC AGC  1259
Cys Ala Asp Gly Ile Val Asp Cys Arg Glu Lys Ser Leu Thr Ser
                305                 310                 315

GTG CCC GTC ACC TTG CCC GAC GAC ACC ACC GAC GTT CGC CTC GAG  1304
Val Pro Val Thr Leu Pro Asp Asp Thr Thr Asp Val Arg Leu Glu
                320                 325                 330

CAA AAT TTC ATT ACG GAA CTG CCG CCG AAA TCG TTC TCC AGC TTT  1349
Gln Asn Phe Ile Thr Glu Leu Pro Pro Lys Ser Phe Ser Ser Phe
                335                 340                 345

CGA CGA CTG CGA CGC ATC GAC CTG TCC AAC AAC AAC ATA TCC CGG  1394
Arg Arg Leu Arg Arg Ile Asp Leu Ser Asn Asn Asn Ile Ser Arg
                350                 355                 360

ATT GCC CAC GAT GCA CTA AGC GGC CTA AAG CAG TTA ACC ACT CTC  1439
Ile Ala His Asp Ala Leu Ser Gly Leu Lys Gln Leu Thr Thr Leu
                365                 370                 375

GTG CTG TAC GGC AAT AAA ATA AAG GAT TTA CCC TCG GGC GTG TTC  1484
Val Leu Tyr Gly Asn Lys Ile Lys Asp Leu Pro Ser Gly Val Phe
                380                 385                 390

AAA GGA CTC GGC TCG CTC AGG CTG CTG CTG CTG AAC GCC AAC GAG  1529
Lys Gly Leu Gly Ser Leu Arg Leu Leu Leu Leu Asn Ala Asn Glu
                395                 400                 405

ATC TCG TGC ATA CGC AAG GAT GCC TTT CGC GAC CTG CAC AGT TTG  1574
Ile Ser Cys Ile Arg Lys Asp Ala Phe Arg Asp Leu His Ser Leu
                410                 415                 420

AGC CTG CTC TCC CTG TAC GAC AAC AAC ATC CAG TCG CTG GCT AAT  1619
Ser Leu Leu Ser Leu Tyr Asp Asn Asn Ile Gln Ser Leu Ala Asn
                425                 430                 435

GGC ACA TTC GAC GCC ATG AAG AGC ATG AAA ACG GTA CAT CTG GCC  1664
Gly Thr Phe Asp Ala Met Lys Ser Met Lys Thr Val His Leu Ala
                440                 445                 450
```

```
AAG AAT CCT TTC ATC TGC GAC TGC AAT CTG CGC TGG CTG GCC GAC  1709
Lys Asn Pro Phe Ile Cys Asp Cys Asn Leu Arg Trp Leu Ala Asp
                455                 460                 465

TAT TTG CAC AAA AAT CCC ATA GAG ACG AGT GGA GCC CGC TGC GAG  1754
Tyr Leu His Lys Asn Pro Ile Glu Thr Ser Gly Ala Arg Cys Glu
                470                 475                 480

TCA CCG AAG CGG ATG CAT CGT CGT CGG ATT GAA TCG CTG CGC GAG  1799
Ser Pro Lys Arg Met His Arg Arg Arg Ile Glu Ser Leu Arg Glu
                485                 490                 495

GAG AAA TTC AAA TGC TCC TGG GGT GAA TTG CGG ATG AAG CTG TCC  1844
Glu Lys Phe Lys Cys Ser Trp Gly Glu Leu Arg Met Lys Leu Ser
                500                 505                 510

GGC GAG TGC CGC ATG GAC TCC GAC TGT CCG GCC ATG TGC CAC TGC  1889
Gly Glu Cys Arg Met Asp Ser Asp Cys Pro Ala Met Cys His Cys
                515                 520                 525

GAG GGC ACC ACC GTG GAT TGC ACG GGC CGG CGG CTG AAG GAG ATT  1934
Glu Gly Thr Thr Val Asp Cys Thr Gly Arg Arg Leu Lys Glu Ile
                530                 535                 540

CCG CGC GAC ATT CCC CTG CAC ACA ACT GAG CTT TTG CTC AAC GAC  1979
Pro Arg Asp Ile Pro Leu His Thr Thr Glu Leu Leu Leu Asn Asp
                545                 550                 555

AAC GAA CTG GGA CGC ATC AGT TCC GAT GGC CTC TTT GGT CGC CTG  2024
Asn Glu Leu Gly Arg Ile Ser Ser Asp Gly Leu Phe Gly Arg Leu
                560                 565                 570

CCG CAC TTG GTG AAG CTG GAA TTG AAG CGC AAC CAG CTG ACC GGC  2069
Pro His Leu Val Lys Leu Glu Leu Lys Arg Asn Gln Leu Thr Gly
                575                 580                 585

ATC GAG CCC AAC GCC TTC GAG GGA GCA TCC CAC ATC CAG GAG TTG  2114
Ile Glu Pro Asn Ala Phe Glu Gly Ala Ser His Ile Gln Glu Leu
                590                 595                 600

CAG CTG GGC GAG AAC AAG ATC AAG GAG ATA TCG AAC AAG ATG TTC  2159
Gln Leu Gly Glu Asn Lys Ile Lys Glu Ile Ser Asn Lys Met Phe
                605                 610                 615

CTG GGA CTG CAC CAA CTA AAA ACG CTC AAT CTG TAC GAC AAT CAA  2204
Leu Gly Leu His Gln Leu Lys Thr Leu Asn Leu Tyr Asp Asn Gln
                620                 625                 630

ATC TCA TGC GTT ATG CCC GGT TCC TTT GAG CAT CTC AAC TCT CTG  2249
Ile Ser Cys Val Met Pro Gly Ser Phe Glu His Leu Asn Ser Leu
                635                 640                 645

ACG TCG CTG AAC CTC GCA TCG AAT CCA TTC AAT TGC AAT TGT CAT  2294
Thr Ser Leu Asn Leu Ala Ser Asn Pro Phe Asn Cys Asn Cys His
                650                 655                 660
```

```
TTG GCC TGG TTC GCG GAA TGC GTG CGC AAA AAA TCA CTG AAC GGC  2339
Leu Ala Trp Phe Ala Glu Cys Val Arg Lys Lys Ser Leu Asn Gly
                665                 670                 675

GGA GCG GCA CGT TGT GGA GCC CCG TCG AAG GTA CGT GAC GTG CAG  2384
Gly Ala Ala Arg Cys Gly Ala Pro Ser Lys Val Arg Asp Val Gln
                680                 685                 690

ATC AAG GAC TTG CCC CAC TCG GAA TTC AAG TGT AGC AGC GAG AAC  2429
Ile Lys Asp Leu Pro His Ser Glu Phe Lys Cys Ser Ser Glu Asn
                695                 700                 705

AGC GAG GGC TGC CTG GGC GAT GGC TAC TGT CCG CCA TCC TGC ACC  2474
Ser Glu Gly Cys Leu Gly Asp Gly Tyr Cys Pro Pro Ser Cys Thr
                710                 715                 720

TGC ACC GGC ACC GTG GTC GCC TGT TCG CGT AAC CAG CTG AAG GAG  2519
Cys Thr Gly Thr Val Val Ala Cys Ser Arg Asn Gln Leu Lys Glu
                725                 730                 735

ATA CCG CGA GGC ATT CCC GCC GAA ACA TCG GAG CTG TAT CTG GAG  2564
Ile Pro Arg Gly Ile Pro Ala Glu Thr Ser Glu Leu Tyr Leu Glu
                740                 745                 750

TCC AAT GAG ATC GAG CAG ATT CAC TAC GAA CGC ATA CGC CAT TTG  2609
Ser Asn Glu Ile Glu Gln Ile His Tyr Glu Arg Ile Arg His Leu
                755                 760                 765

CGC TCC CTT ACC CGA CTC GAT CTC AGC AAC AAC CAG ATC ACC ATT  2654
Arg Ser Leu Thr Arg Leu Asp Leu Ser Asn Asn Gln Ile Thr Ile
                770                 775                 780

CTT TCC AAC TAC ACC TTT GCC AAT CTG ACC AAG CTG TCC ACG CTC  2699
Leu Ser Asn Tyr Thr Phe Ala Asn Leu Thr Lys Leu Ser Thr Leu
                785                 790                 795

ATC ATC TCA TAC AAC AAG CTG CAG TGT CTG CAG CGG CAT GCG TTG  2744
Ile Ile Ser Tyr Asn Lys Leu Gln Cys Leu Gln Arg His Ala Leu
                800                 805                 810

TCT GGC CTG AAT AAC CTG CGC GTC GTT TCG CTG CAC GGT AAC CGC  2789
Ser Gly Leu Asn Asn Leu Arg Val Val Ser Leu His Gly Asn Arg
                815                 820                 825

ATC TCG ATG CTG CCG GAA GGC TCC TTC GAG GAC CTC AAG TCG TTG  2834
Ile Ser Met Leu Pro Glu Gly Ser Phe Glu Asp Leu Lys Ser Leu
                830                 835                 840

ACC CAC ATC GCA CTA GGC AGC AAT CCC TTG TAC TGC GAC TGC GGT  2879
Thr His Ile Ala Leu Gly Ser Asn Pro Leu Tyr Cys Asp Cys Gly
                845                 850                 855

CTA AAG TGG TTC TCC GAT TGG ATC AAG CTG GAC TAC GTG GAA CCG  2924
Leu Lys Trp Phe Ser Asp Trp Ile Lys Leu Asp Tyr Val Glu Pro
                860                 865                 870
```

```
GGA ATT GCA CGT TGC GCC GAA CCG GAA CAG ATG AAG GAT AAG CTG  2969
Gly Ile Ala Arg Cys Ala Glu Pro Glu Gln Met Lys Asp Lys Leu
                875                 880                 885

ATC CTG TCC ACA CCC TCG TCG AGC TTC GTT TGC CGC GGC CGC GTG  3014
Ile Leu Ser Thr Pro Ser Ser Ser Phe Val Cys Arg Gly Arg Val
                890                 895                 900

CGC AAT GAT ATT CTG GCC AAG TGC AAC GCC TGT TTC GAG CAG CCA  3059
Arg Asn Asp Ile Leu Ala Lys Cys Asn Ala Cys Phe Glu Gln Pro
                905                 910                 915

TGC CAG AAT CAG GCG CAG TGT GTG GCC CTT CCG CAG CGA GAG TAC  3104
Cys Gln Asn Gln Ala Gln Cys Val Ala Leu Pro Gln Arg Glu Tyr
                920                 925                 930

CAG TGC CTC TGC CAG CCG GGC TAT CAT GGG AAA CAC TGT GAG TTT  3149
Gln Cys Leu Cys Gln Pro Gly Tyr His Gly Lys His Cys Glu Phe
                935                 940                 945

ATG ATC GAT GCT TGC TAC GGA AAT CCG TGC CGC AAC AAT GCC ACC  3194
Met Ile Asp Ala Cys Tyr Gly Asn Pro Cys Arg Asn Asn Ala Thr
                950                 955                 960

TGC ACG GTG CTG GAG GAG GGT CGA TTC AGC TGT CAG TGC GCT CCG  3239
Cys Thr Val Leu Glu Glu Gly Arg Phe Ser Cys Gln Cys Ala Pro
                965                 970                 975

GGA TAC ACA GGT GCC CGC TGC GAG ACG AAT ATC GAC GAT TGC CTG  3284
Gly Tyr Thr Gly Ala Arg Cys Glu Thr Asn Ile Asp Asp Cys Leu
                980                 985                 990

GGC GAG ATC AAG TGC CAG AAC AAT GCC ACC TGC ATC GAC GGA GTG  3329
Gly Glu Ile Lys Cys Gln Asn Asn Ala Thr Cys Ile Asp Gly Val
                995                 1000                1005

GAG TCG TAC AAA TGT GAG TGC CAG CCG GGA TTC AGT GGC GAG TTC  3374
Glu Ser Tyr Lys Cys Glu Cys Gln Pro Gly Phe Ser Gly Glu Phe
                1010                1015                1020

TGC GAC ACC AAA ATC CAG TTC TGC AGT CCG GAG TTC AAT CCC TGC  3419
Cys Asp Thr Lys Ile Gln Phe Cys Ser Pro Glu Phe Asn Pro Cys
                1025                1030                1035

GCG AAT GGA GCC AAG TGC ATG GAC CAC TTT ACC CAC TAC AGC TGC  3464
Ala Asn Gly Ala Lys Cys Met Asp His Phe Thr His Tyr Ser Cys
                1040                1045                1050

GAT TGT CAG GCA GGT TTC CAT GGC ACC AAC TGC ACG GAC AAT ATT  3509
Asp Cys Gln Ala Gly Phe His Gly Thr Asn Cys Thr Asp Asn Ile
                1055                1060                1065

GAC GAC TGC CAG AAC CAC ATG TGC CAG AAC GGT GGA ACG TGC GTG  3554
Asp Asp Cys Gln Asn His Met Cys Gln Asn Gly Gly Thr Cys Val
                1070                1075                1080
```

```
GAC GGC ATC AAC GAC TAC CAA TGC CGC TGT CCA GAC GAC TAT ACG  3599
Asp Gly Ile Asn Asp Tyr Gln Cys Arg Cys Pro Asp Asp Tyr Thr
                1085                1090                1095

GGC AAG TAC TGT GAA GGC CAC AAC ATG ATC TCG ATG ATG TAT CCA  3644
Gly Lys Tyr Cys Glu Gly His Asn Met Ile Ser Met Met Tyr Pro
                1100                1105                1110

CAG ACG TCG CCT TGT CAA AAC CAC GAG TGC AAG CAC GGT GTC TGC  3689
Gln Thr Ser Pro Cys Gln Asn His Glu Cys Lys His Gly Val Cys
                1115                1120                1125

TTC CAA CCG AAC GCT CAG GGC AGC GAC TAC CTA TGC AGG TGT CAT  3734
Phe Gln Pro Asn Ala Gln Gly Ser Asp Tyr Leu Cys Arg Cys His
                1130                1135                1140

CCG GGT TAC ACT GGA AAG TGG TGC GAG TAC CTC ACC AGC ATT AGC  3779
Pro Gly Tyr Thr Gly Lys Trp Cys Glu Tyr Leu Thr Ser Ile Ser
                1145                1150                1155

TTC GTC CAC AAC AAC TCG TTT GTG GAA CTG GAG CCA CTG CGA ACA  3824
Phe Val His Asn Asn Ser Phe Val Glu Leu Glu Pro Leu Arg Thr
                1160                1165                1170

CGT CCG GAG GCG AAC GTG ACG ATA GTC TTC AGC AGC GCG GAG CAG  3869
Arg Pro Glu Ala Asn Val Thr Ile Val Phe Ser Ser Ala Glu Gln
                1175                1180                1185

AAT GGA ATT CTC ATG TAC GAC GGC CAG GAT GCA CAT CTC GCA GTG  3914
Asn Gly Ile Leu Met Tyr Asp Gly Gln Asp Ala His Leu Ala Val
                1190                1195                1200

GAG CTG TTT AAT GGG CGT ATT CGG GTT AGC TAC GAT GTG GGT AAT  3959
Glu Leu Phe Asn Gly Arg Ile Arg Val Ser Tyr Asp Val Gly Asn
                1205                1210                1215

CAC CCT GTG TCC ACG ATG TAC AGC TTT GAA ATG GTG GCC GAT GGA  4004
His Pro Val Ser Thr Met Tyr Ser Phe Glu Met Val Ala Asp Gly
                1220                1225                1230

AAG TAC CAT GCC GTG GAG CTT CTG GCC ATC AAG AAG AAT TTC ACG  4049
Lys Tyr His Ala Val Glu Leu Leu Ala Ile Lys Lys Asn Phe Thr
                1235                1240                1245

CTG CGC GTG GAT CGC GGA TTG GCC CGT TCC ATC ATC AAC GAG GGC  4094
Leu Arg Val Asp Arg Gly Leu Ala Arg Ser Ile Ile Asn Glu Gly
                1250                1255                1260

TCC AAC GAC TAC CTG AAA CTT ACG ACT CCC ATG TTC CTG GGC GGC  4139
Ser Asn Asp Tyr Leu Lys Leu Thr Thr Pro Met Phe Leu Gly Gly
                1265                1270                1275

CTA CCA GTG GAT CCT GCA CAG CAG GCA TAC AAG AAC TGG CAA ATA  4184
Leu Pro Val Asp Pro Ala Gln Gln Ala Tyr Lys Asn Trp Gln Ile
                1280                1285                1290
```

```
CGC AAC CTT ACC AGC TTT AAG GGC TGC ATG AAG GAG GTG TGG ATC  4229
Arg Asn Leu Thr Ser Phe Lys Gly Cys Met Lys Glu Val Trp Ile
                1295                1300                1305

AAT CAT AAG CTG GTC GAC TTT GGC AAT GCC CAG CGC CAG CAA AAG  4274
Asn His Lys Leu Val Asp Phe Gly Asn Ala Gln Arg Gln Gln Lys
                1310                1315                1320

ATC ACA CCA GGA TGT GCC CTG CTC GAA GGA GAG CAG CAA GAG GAG  4319
Ile Thr Pro Gly Cys Ala Leu Leu Glu Gly Glu Gln Gln Glu Glu
                1325                1330                1335

GAA GAC GAC GAG CAG GAT TTC ATG GAC GAG ACA CCG CAC ATC AAA  4364
Glu Asp Asp Glu Gln Asp Phe Met Asp Glu Thr Pro His Ile Lys
                1340                1345                1350

GAG GAG CCG GTG GAT CCT TGC CTG GAG AAC AAA TGC CGT CGG GGC  4409
Glu Glu Pro Val Asp Pro Cys Leu Glu Asn Lys Cys Arg Arg Gly
                1355                1360                1365

AGT CGC TGT GTG CCG AAT TCC AAT GCC AGG GAC GGC TAC CAG TGC  4454
Ser Arg Cys Val Pro Asn Ser Asn Ala Arg Asp Gly Tyr Gln Cys
                1370                1375                1380

AAG TGC AAG CAC GGC CAG CGC GGC CGC TAC TGC GAT CAA GGT GAG  4499
Lys Cys Lys His Gly Gln Arg Gly Arg Tyr Cys Asp Gln Gly Glu
                1385                1390                1395

GGC AGC ACT GAG CCC CCA ACA GTC ACC GCG GCG TCC ACC TGT CGC  4544
Gly Ser Thr Glu Pro Pro Thr Val Thr Ala Ala Ser Thr Cys Arg
                1400                1405                1410

AAG GAG CAG GTG CGC GAG TAC TAC ACG GAG AAC GAC TGT CGC TCG  4589
Lys Glu Gln Val Arg Glu Tyr Tyr Thr Glu Asn Asp Cys Arg Ser
                1415                1420                1425

AGG CAG CCG TTG AAG TAC GCC AAG TGC GTG GGC GGC TGC GGC AAC  4634
Arg Gln Pro Leu Lys Tyr Ala Lys Cys Val Gly Gly Cys Gly Asn
                1430                1435                1440

CAG TGC TGC GCG GCC AAA ATT GTG AGA CGG CGC AAG GTG CGC ATG  4679
Gln Cys Cys Ala Ala Lys Ile Val Arg Arg Arg Lys Val Arg Met
                1445                1450                1455

GTG TGC AGC AAC AAC CGC AAG TAC ATC AAG AAC TTG GAC ATC GTG  4724
Val Cys Ser Asn Asn Arg Lys Tyr Ile Lys Asn Leu Asp Ile Val
                1460                1465                1470

CGC AAG TGC GGA TGC ACC AAG AAA TGC TAC  4754
Arg Lys Cys Gly Cys Thr Lys Lys Cys Tyr
                1475                1480

                TGACTG AAAGATGCGA CTACCCAATT GCTCGAACGG AGCAATAGCA  4800

GCTTAGATGT TAGTTTAGGA ACAGGTTTAA ATCTAACTTA TAGTAGTAGT AATAGTAACG  4860

ATAGTCTTAG CCATAGCACT AGGGATAGCA CGGATGTTAG GGGGACGAAG GATGAAGTGG  4920
```

AGGAGAGTGC TGACGCGGGG GAGACAACGG AGGAGGTGAG GGACGAGAGT GATGATTACC 4980

CCACGGACGG CATGCAGTCC GATCTCTACG ACGACACCAT CGATGACGAC GATGACGATG 5040

GTCTCGATGA TGATTATGCA GACGAGGAGG ATGGGGAGGA GGATCCAGAG CAGCTTCCCG 5100

ATCCCAAGGG TCTGGTAAGT GTGCCAGATG AAGAGGAAGA CATGGGTTAC GACGAGGATG 5160

ACGAACGTAT CGCCATGGAG CGGCCAAGAA CGGTTAGACC CAGGCCCGAC GAGGAGCATT 5220

TCCTTAACGA AGAGGGCAGT GGTTTCGGTG GCTTTCGATC GCGATTCCGG CCGAGCAATA 5280

GCTTCCGCGA GACTCAGCTG GAGAATATCC GCAAGAAGCT GCTCACAGAA GCACAGGCGG 5340

CTCCGGAAAC GGCTGTTGCG GTGGCCGTGC CGAGCACTGC GATAGATCTG CGCGAGAGCA 5400

GCGGCCACTT TGCCAACGAT GACGAGGATG GCGAGGACGG CGATGACGGC GTCGATGACG 5460

AATTCGCCGA CACGGGGGAG AACCAGGGGC GCGGCTTCTT TGGCTCCCAG CAGCAACAGC 5520

GCAAGAACGG TCCGTATCAC CGCAAGAACG GCAACGATGC CATCAAAATC ATCTCCACGC 5580

CGCTGGGCAA GGTGAGCATT GTGTACCAGC AGACGGACAA GGACCAAAGT CCGGACAAGG 5640

ATGCGCAACA GCAGCAGCAA AANAAGCCGG CGCTCACCGA CTTCGACGCC CTGTCACCGG 5700

ACCCCGAGAG CAGTCATCGC TTTCCGTCGC CCCACCCCAA GATTACACCT GTTCTGACGC 5760

CCGATGGCAA GGTGGCGCTG CTCTATCGCG GAGACTCGGA GAGCTCCAAG TACGAGCCCA 5820

TACGCAACCT GACGCACAAG TTTTCGGGAC AGCCGGCCAA GGAGTCAAAG CCTAAAACCG 5880

AAGATTTCTT CTCGGCGGAG GACTCTNTCT ACACGGACAG CGAGGATACC GAGGACAGTA 5940

AAGGTGAAAC TAGTGCTGGA AAGTCCCCAC CAGTGGCCAG CACGCCCAAG CNACTGCAGC 6000

CAGAGATTTT GGAGCCTCCA GATAACGTCC AGCCAGGAGG GTTATTTATA ATTCGGCCCA 6060

CCTCGGACTC GCTTCTGCCG ATGATCAACA GGCCACTGTC CGAGGTTCTG GGCATCAAGA 6120

AGAACCAGTT CCAGGAGACC CGGGTGCGTG ACCAATTGCC CACGCAACAG CCTCCGCCTC 6180

CACTGCCGGA GGCCACGTCC CGTAGTCCCG CTTCCGGACA TCAGTTCCTA GCCAAGGTGA 6240

ACCTGGCTGA GTTCCCGACA TCCGGAAGGA CGCTCCAGAC GCCGCTGATC CCCAGCACCC 6300

ACGACTTTGA CTTCAGCCGT GACAACACGA TGCTGGACGA GCGGTCGCGG GTGCGTGAGT 6360

TGGAAAAGCA GCGGGAGCGG GACAAGGAGC ACAACGAGGC CACCAGCAAG GGAGCCACCG 6420

AGGCACACAC CATAGCCATA CGAGCAGCTG CTCTCCAAGA CGGAAGTCAT CAATCTGGCC 6480

ATTGTGCCAC AGTTTGACGA GGACTTGGAG CGTCTACAGC GGCTGCAGGA GAATGGTGGT 6540

CGGCGCCACC ATAGGGCGCG ACATCGTCAT CGGCAACAGT CGGAGGAGGA ACTGTCTGGC 6600

ATCCATTGCA TCATGCAGGT CATGATGGCG TGGCCGCCGT GTCGACGGTC TTCGGCATGC 6660

TGGGCACCTT CTTCAAGCAA CGCATCCTCG ATCANTGCGC ATGATGCACT GGTAGTACAA 6720

CAAGGGGGAT TCGATTTTCG GTGTCGAATG CCCCACATTC CCCCACCCAA TCCCCGTTTC 6780

CTGCATCGTC ATGATCATCA TGCAGTCTAT GGAGGCTGGC TAGCTCNCTT CGGGATGCGA 6840

GGTCCTTCTT CTACTACTAG CACTCATATA CTCGAATATA TACTCGTACT CGTACCATAT 6900

GCCATATGCC ATATATTAAT CGCATAATCT ATGTAACACA GCGGCATCGA TTTGCTTTCG 6960

NCCCCTTCCG CTTCCTTATA TATATTTATA TATACATTTA TATATCTATC TTATCCTTCG 7020

GCATTGTGCT CGNAAATGCG GACACTTCCT TGCTACACAT GTATTCTATA TATTCGCATA 7080

TATCAATTTA CTAGTGCAAG CTACCCAGGC GATGTACATA ATAACAGAAC CATAATACGG 7140

CGATCGATCG ATCCGGATAT CTATGTATTT ATGTGAGACG CAACTGTAGC CCCTGCCACA 7200

CTCCGCCGGA AATTAGCTGA CCCACGTCAC TTCTTCCCCG TTCCTTCTTT CGCACGTTGT 7260

GTGCCCTATT TTCCCCTCCC ATGTAGGACA TTCTNAATCA CATCAGCGTA TACAAGCTTT 7320

AAGCAAGTCA TTGCATGTGC CACGCCCCCT CGAACTGAAC TGAACTCAAC GCCCAACCCG 7380

CCCAGCATGA GTCCCGCAGC AAATACATTC CCTCCATCCC CCCACCATCC GTCATCCGCC 7440

AGAGGAACCA GTCGAGGAGC TCGGAAATTG AACGTAGAAC AGATTCCTTT TGTAGATAGA 7500

AAACGAAACA GTTGTTGGGA GATGGTAACC AGACGAATGT CGAATGACAA ACGATAAATA 7560

ATGATAAACT AAACTAAAGT TCTAAACAAA ACAAAACACA GTAAAATCGC ACAGAAGCGC 7620

ACGCATTACA AAATACAAAA ACCTGCAACG GTCGTTTTAA AACGCTCCGT TCAGTTGTCT 7680

CAAAGAAACG AGTAAACGAT AATAAGTGCA TAACGAAAAC CTTCTTTAGT CTAGTTGCTC 7740

AAAGGATAAA GTATTTTGAT AGAACCGGAA AGGATCGAGA AACAAACCAA TAAACCACAA 7800

AGAATCGAAT TGAATCGTAA CGAAACAAAG GCCCCAAACG TGTAACGAAT TTCCAACAAA 7860

TTGTTGCAAG TGTTTTCCTA GAATTAGTCC TAATTAAACT AAATGTGTGC AAATCGAAGC 7920

GTAACTAATA TTACAATTAA TCTAAACTAA TTGAGAAACC ATAAACCTAA ACATTAAATC 7980

GGAAAACAAC ATCTAAGCTG GGTAGTCGCA TGTAAATCTC TAACAATTAA CAATTACCGG 8040

CCTAAGTTAG ACCTAAAAAT CGAAACAAAA TCGAATCGCA TTTAAAGAAA TCTACATAAT 8100

AATAATAATT TATACTAATC TATATATACT TATATGTATG CTGTATGTAT GTATGACCCT 8160

ATATGTATAT GTAAAATGTT TTTGACTATT TTTCACTATT TATATTCATA TATATTATAT 8220

ATGCATATAC GATACATGTG TAATAGCCCC TTTTGGTCAT TTTAGTTGTC TTTTATATAT 8280

ATTTAATACG TGTATTATTT TTATTAATTA TTCAAGTATA ACTATGCGCA CCAATTAAAC 8340

GCATACCTTA TGTATAACCT ATTGACAAAA AAAAAAAA 8378

(2) INFORMATION FOR SEQ ID NO: 2:
(i) SEQUENCE CHARACTERESITICS:
(A) LENGTH: 1480 amino acids
(B) TYPE: amino acids
(D) TOPOLOGY: Linear
(ii) MOLECULE TYPE: protein
(ix) FEATURE:
(A) NAME/KEY: signal sequence
(B) LOCATION: 1 to 36
(C) IDENTIFICATION METHOD: similarity to other signal sequences.
(D) OTHER INFORMATION: Directs Export

(A) NAME/KEY: Four Flank-LRR-Flank domains
(B) LOCATION: 37 to 910
(C) IDENTIFICATION METHOD: Array of Flank-LRR-Flank domains defined herein.
(D) OTHER INFORMATION: mediates adhesive events

(A) NAME/KEY: Tandem EGF-like repeats
(B) LOCATION: 911 to 1150
(C) IDENTIFICATION METHOD: similarity to tandem EGF-like repeats
(D) OTHER INFORMATION: protein-protein interactions

(A) NAME/KEY: 7th EGF-like repeat
(B) LOCATION: 1353 to 1393
(C) IDENTIFICATION METHOD: similarity to epidermal growth factor
(D) OTHER INFORMATION: Involvement in receptor-ligand interactions

(A) NAME/KEY: Alternative splice segment
(B) LOCATION: 1394 to 1404
(C) IDENTIFICATION METHOD: experimental
(D) OTHER INFORMATION: developmentally regulated

(A) NAME/KEY: COOH-terminal region
(B) LOCATION: 1405 to 1480
(C) IDENTIFICATION METHOD: experimental

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```
Met Ala Ala Pro Ser Arg Thr Thr Leu Met Pro Pro Pro Phe Arg
                5                  10                  15

Leu Gln Leu Arg Leu Leu Ile Leu Pro Ile Leu Leu Leu Leu Arg
            20                  25                  30

His Asp Ala Val His Ala Glu Pro Tyr Ser Gly Gly Phe Gly Ser
        35                  40                  45

Ser Ala Val Ser Ser Gly Gly Leu Gly Ser Val Gly Ile His Ile
    50                  55                  60

Pro Gly Gly Gly Val Gly Val Ile Thr Glu Ala Arg Cys Pro Arg
65                  70                  75
```

Val Cys Ser Cys Thr Gly Leu Asn Val Asp Cys Ser His Arg Gly
80 85 90

Leu Thr Ser Val Pro Arg Lys Ile Ser Ala Asp Val Glu Arg Leu
95 100 105

Glu Leu Gln Gly Asn Asn Leu Thr Val Ile Tyr Glu Thr Asp Phe
110 115 120

Gln Arg Leu Thr Lys Leu Arg Met Leu Gln Leu Thr Asp Asn Gln
125 130 135

Ile His Thr Ile Glu Arg Asn Ser Phe Gln Asp Leu Val Ser Leu
140 145 150

Glu Arg Leu Asp Ile Ser Asn Asn Val Ile Thr Thr Val Gly Arg
155 160 165

Arg Val Phe Lys Gly Ala Gln Ser Leu Arg Ser Leu Gln Leu Asp
170 175 180

Asn Asn Gln Ile Thr Cys Leu Asp Glu His Ala Phe Lys Gly Leu
185 190 195

Val Glu Leu Glu Ile Leu Thr Leu Asn Asn Asn Asn Leu Thr Ser
200 205 210

Leu Pro His Asn Ile Phe Gly Gly Leu Gly Arg Leu Arg Ala Leu
215 220 225

Arg Leu Ser Asp Asn Pro Phe Ala Cys Asp Cys His Leu Ser Trp
230 235 240

Leu Ser Arg Phe Leu Arg Ser Ala Thr Arg Leu Ala Pro Tyr Thr
245 250 255

Arg Cys Gln Ser Pro Ser Gln Leu Lys Gly Gln Asn Val Ala Asp
260 265 270

Leu His Asp Gln Glu Phe Lys Cys Ser Gly Leu Thr Glu His Ala
275 280 285

Pro Met Glu Cys Gly Ala Glu Asn Ser Cys Pro His Pro Cys Arg
290 295 300

Cys Ala Asp Gly Ile Val Asp Cys Arg Glu Lys Ser Leu Thr Ser
305 310 315

Val Pro Val Thr Leu Pro Asp Asp Thr Thr Asp Val Arg Leu Glu
320 325 330

Gln Asn Phe Ile Thr Glu Leu Pro Pro Lys Ser Phe Ser Ser Phe
335 340 345

Arg Arg Leu Arg Arg Ile Asp Leu Ser Asn Asn Asn Ile Ser Arg
350 355 360

Ile Ala His Asp Ala Leu Ser Gly Leu Lys Gln Leu Thr Thr Leu
365 370 375

Val Leu Tyr Gly Asn Lys Ile Lys Asp Leu Pro Ser Gly Val Phe
380 385 390

Lys Gly Leu Gly Ser Leu Arg Leu Leu Leu Leu Asn Ala Asn Glu
395 400 405

Ile Ser Cys Ile Arg Lys Asp Ala Phe Arg Asp Leu His Ser Leu
410 415 420

Ser Leu Leu Ser Leu Tyr Asp Asn Asn Ile Gln Ser Leu Ala Asn
425 430 435

Gly Thr Phe Asp Ala Met Lys Ser Met Lys Thr Val His Leu Ala
440 445 450

Lys Asn Pro Phe Ile Cys Asp Cys Asn Leu Arg Trp Leu Ala Asp
455 460 465

Tyr Leu His Lys Asn Pro Ile Glu Thr Ser Gly Ala Arg Cys Glu
470 475 480

Ser Pro Lys Arg Met His Arg Arg Arg Ile Glu Ser Leu Arg Glu
485 490 495

Glu Lys Phe Lys Cys Ser Trp Gly Glu Leu Arg Met Lys Leu Ser
500 505 510

Gly Glu Cys Arg Met Asp Ser Asp Cys Pro Ala Met Cys His Cys
515 520 525

Glu Gly Thr Thr Val Asp Cys Thr Gly Arg Arg Leu Lys Glu Ile
530 535 540

Pro Arg Asp Ile Pro Leu His Thr Thr Glu Leu Leu Leu Asn Asp
545 550 555

Asn Glu Leu Gly Arg Ile Ser Ser Asp Gly Leu Phe Gly Arg Leu
560 565 570

Pro His Leu Val Lys Leu Glu Leu Lys Arg Asn Gln Leu Thr Gly
575 580 585

Ile Glu Pro Asn Ala Phe Glu Gly Ala Ser His Ile Gln Glu Leu
590 595 600

Gln Leu Gly Glu Asn Lys Ile Lys Glu Ile Ser Asn Lys Met Phe
605 610 615

Leu Gly Leu His Gln Leu Lys Thr Leu Asn Leu Tyr Asp Asn Gln
620 625 630

Ile Ser Cys Val Met Pro Gly Ser Phe Glu His Leu Asn Ser Leu
635 640 645

Thr Ser Leu Asn Leu Ala Ser Asn Pro Phe Asn Cys Asn Cys His
650 655 660

Leu Ala Trp Phe Ala Glu Cys Val Arg Lys Lys Ser Leu Asn Gly
665 670 675

Gly Ala Ala Arg Cys Gly Ala Pro Ser Lys Val Arg Asp Val Gln
680 685 690

Ile Lys Asp Leu Pro His Ser Glu Phe Lys Cys Ser Ser Glu Asn
695 700 705

Ser Glu Gly Cys Leu Gly Asp Gly Tyr Cys Pro Pro Ser Cys Thr
710 715 720

Cys Thr Gly Thr Val Val Ala Cys Ser Arg Asn Gln Leu Lys Glu
725 730 735

Ile Pro Arg Gly Ile Pro Ala Glu Thr Ser Glu Leu Tyr Leu Glu
740 745 750

Ser Asn Glu Ile Glu Gln Ile His Tyr Glu Arg Ile Arg His Leu
755 760 765

Arg Ser Leu Thr Arg Leu Asp Leu Ser Asn Asn Gln Ile Thr Ile
770 775 780

Leu Ser Asn Tyr Thr Phe Ala Asn Leu Thr Lys Leu Ser Thr Leu
785 790 795

Ile Ile Ser Tyr Asn Lys Leu Gln Cys Leu Gln Arg His Ala Leu
800 805 810

Ser Gly Leu Asn Asn Leu Arg Val Val Ser Leu His Gly Asn Arg
815 820 825

Ile Ser Met Leu Pro Glu Gly Ser Phe Glu Asp Leu Lys Ser Leu
830 835 840

Thr His Ile Ala Leu Gly Ser Asn Pro Leu Tyr Cys Asp Cys Gly
845 850 855

Leu Lys Trp Phe Ser Asp Trp Ile Lys Leu Asp Tyr Val Glu Pro
860 865 870

Gly Ile Ala Arg Cys Ala Glu Pro Glu Gln Met Lys Asp Lys Leu
875 880 885

Ile Leu Ser Thr Pro Ser Ser Ser Phe Val Cys Arg Gly Arg Val
890 895 900

Arg Asn Asp Ile Leu Ala Lys Cys Asn Ala Cys Phe Glu Gln Pro
905 910 915

Cys Gln Asn Gln Ala Gln Cys Val Ala Leu Pro Gln Arg Glu Tyr
920 925 930

Gln Cys Leu Cys Gln Pro Gly Tyr His Gly Lys His Cys Glu Phe
935 940 945

Met Ile Asp Ala Cys Tyr Gly Asn Pro Cys Arg Asn Asn Ala Thr
950 955 960

Cys Thr Val Leu Glu Glu Gly Arg Phe Ser Cys Gln Cys Ala Pro
965 970 975

Gly Tyr Thr Gly Ala Arg Cys Glu Thr Asn Ile Asp Asp Cys Leu
980 985 990

Gly Glu Ile Lys Cys Gln Asn Asn Ala Thr Cys Ile Asp Gly Val
995 1000 1005

Glu Ser Tyr Lys Cys Glu Cys Gln Pro Gly Phe Ser Gly Glu Phe
1010 1015 1020

Cys Asp Thr Lys Ile Gln Phe Cys Ser Pro Glu Phe Asn Pro Cys
1025 1030 1035

Ala Asn Gly Ala Lys Cys Met Asp His Phe Thr His Tyr Ser Cys
1040 1045 1050

Asp Cys Gln Ala Gly Phe His Gly Thr Asn Cys Thr Asp Asn Ile
1055 1060 1065

Asp Asp Cys Gln Asn His Met Cys Gln Asn Gly Gly Thr Cys Val
1070 1075 1080

Asp Gly Ile Asn Asp Tyr Gln Cys Arg Cys Pro Asp Asp Tyr Thr
1085 1090 1095

Gly Lys Tyr Cys Glu Gly His Asn Met Ile Ser Met Met Tyr Pro
1100 1105 1110

Gln Thr Ser Pro Cys Gln Asn His Glu Cys Lys His Gly Val Cys
1115 1120 1125

Phe Gln Pro Asn Ala Gln Gly Ser Asp Tyr Leu Cys Arg Cys His
1130 1135 1140

Pro Gly Tyr Thr Gly Lys Trp Cys Glu Tyr Leu Thr Ser Ile Ser
1145 1150 1155

Phe Val His Asn Asn Ser Phe Val Glu Leu Glu Pro Leu Arg Thr
1160 1165 1170

Arg Pro Glu Ala Asn Val Thr Ile Val Phe Ser Ser Ala Glu Gln
1175 1180 1185

Asn Gly Ile Leu Met Tyr Asp Gly Gln Asp Ala His Leu Ala Val
1190 1195 1200

Glu Leu Phe Asn Gly Arg Ile Arg Val Ser Tyr Asp Val Gly Asn
1205 1210 1215

His Pro Val Ser Thr Met Tyr Ser Phe Glu Met Val Ala Asp Gly
1220 1225 1230

Lys Tyr His Ala Val Glu Leu Leu Ala Ile Lys Lys Asn Phe Thr
1235 1240 1245

Leu Arg Val Asp Arg Gly Leu Ala Arg Ser Ile Ile Asn Glu Gly
1250 1255 1260

Ser Asn Asp Tyr Leu Lys Leu Thr Thr Pro Met Phe Leu Gly Gly
1265 1270 1275

Leu Pro Val Asp Pro Ala Gln Gln Ala Tyr Lys Asn Trp Gln Ile
1280 1285 1290

Arg Asn Leu Thr Ser Phe Lys Gly Cys Met Lys Glu Val Trp Ile
1295 1300 1305

Asn His Lys Leu Val Asp Phe Gly Asn Ala Gln Arg Gln Gln Lys
1310 1315 1320

Ile Thr Pro Gly Cys Ala Leu Leu Glu Gly Glu Gln Gln Glu Glu
1325 1330 1335

Glu Asp Asp Glu Gln Asp Phe Met Asp Glu Thr Pro His Ile Lys
1340 1345 1350

Glu Glu Pro Val Asp Pro Cys Leu Glu Asn Lys Cys Arg Arg Gly
1355 1360 1365

Ser Arg Cys Val Pro Asn Ser Asn Ala Arg Asp Gly Tyr Gln Cys
1370 1375 1380

Lys Cys Lys His Gly Gln Arg Gly Arg Tyr Cys Asp Gln Gly Glu
1385 1390 1395

Gly Ser Thr Glu Pro Pro Thr Val Thr Ala Ala Ser Thr Cys Arg
1400 1405 1410

Lys Glu Gln Val Arg Glu Tyr Tyr Thr Glu Asn Asp Cys Arg Ser
1415 1420 1425

Arg Gln Pro Leu Lys Tyr Ala Lys Cys Val Gly Gly Cys Gly Asn
1430 1435 1440

Gln Cys Cys Ala Ala Lys Ile Val Arg Arg Arg Lys Val Arg Met
1445 1450 1455

Val Cys Ser Asn Asn Arg Lys Tyr Ile Lys Asn Leu Asp Ile Val
1460 1465 1470

Arg Lys Cys Gly Cys Thr Lys Lys Cys Tyr
1475 1480

(2) INFORMATION FOR SEQ ID NO: 3:
(i) SEQUENCE CHARACTERESITICS:
(A) LENGTH: 222 amino acids
(B) TYPE: amino acids
(D) TOPOLOGY: Linear
(ii) MOLECULE TYPE: protein
(ix) FEATURE:
(A) NAME/KEY: Flank-LRR-Flank 1
(B) LOCATION: 1 to 222
(C) IDENTIFICATION METHOD: similarity to other Flank-LRR-Flank domains defined herein.
(D) OTHER INFORMATION: mediates adhesive events
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Cys Pro Arg Val Cys Ser Cys Thr Gly Leu Asn Val Asp Cys Ser
5 10 15

His Arg Gly Leu Thr Ser Val Pro Arg Lys Ile Ser Ala Asp Val
20 25 30

Glu Arg Leu Glu Leu Gln Gly Asn Asn Leu Thr Val Ile Tyr Glu
35 40 45

Thr Asp Phe Gln Arg Leu Thr Lys Leu Arg Met Leu Gln Leu Thr
50 55 60

Asp Asn Gln Ile His Thr Ile Glu Arg Asn Ser Phe Gln Asp Leu
65 70 75

Val Ser Leu Glu Arg Leu Asp Ile Ser Asn Asn Val Ile Thr Thr
80 85 90

Val Gly Arg Arg Val Phe Lys Gly Ala Gln Ser Leu Arg Ser Leu
95 100 105

Gln Leu Asp Asn Asn Gln Ile Thr Cys Leu Asp Glu His Ala Phe
110 115 120

Lys Gly Leu Val Glu Leu Glu Ile Leu Thr Leu Asn Asn Asn Asn
125 130 135

Leu Thr Ser Leu Pro His Asn Ile Phe Gly Gly Leu Gly Arg Leu
140 145 150

Arg Ala Leu Arg Leu Ser Asp Asn Pro Phe Ala Cys Asp Cys His
155 160 165

Leu Ser Trp Leu Ser Arg Phe Leu Arg Ser Ala Thr Arg Leu Ala
170 175 180

Pro Tyr Thr Arg Cys Gln Ser Pro Ser Gln Leu Lys Gly Gln Asn
185 190 195

Val Ala Asp Leu His Asp Gln Glu Phe Lys Cys Ser Gly Leu Thr
200 205 210

Glu His Ala Pro Met Glu Cys Gly Ala Glu Asn Ser
215 220

(2) INFORMATION FOR SEQ ID NO: 4:
(i) SEQUENCE CHARACTERESITICS:
(A) LENGTH: 224 amino acids
(B) TYPE: amino acids
(D) TOPOLOGY: Linear
(ii) MOLECULE TYPE: protein
(ix) FEATURE:
(A) NAME/KEY: Flank-LRR-Flank 2
(B) LOCATION: 1 to 224
(C) IDENTIFICATION METHOD: similarity to other Flank-LRR-Flank domains defined herein.
(D) OTHER INFORMATION: mediates adhesive events
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Cys Pro His Pro Cys Arg Cys Ala Asp Gly Ile Val Asp Cys Arg
5 10 15

Glu Lys Ser Leu Thr Ser Val Pro Val Thr Leu Pro Asp Asp Thr
20 25 30

Thr Asp Val Arg Leu Glu Gln Asn Phe Ile Thr Glu Leu Pro Pro
35 40 45

Lys Ser Phe Ser Ser Phe Arg Arg Leu Arg Arg Ile Asp Leu Ser
50 55 60

Asn Asn Asn Ile Ser Arg Ile Ala His Asp Ala Leu Ser Gly Leu
65 70 75

Lys Gln Leu Thr Thr Leu Val Leu Tyr Gly Asn Lys Ile Lys Asp
80 85 90

Leu Pro Ser Gly Val Phe Lys Gly Leu Gly Ser Leu Arg Leu Leu
95 100 105

Leu Leu Asn Ala Asn Glu Ile Ser Cys Ile Arg Lys Asp Ala Phe
110 115 120

Arg Asp Leu His Ser Leu Ser Leu Leu Ser Leu Tyr Asp Asn Asn
125 130 135

Ile Gln Ser Leu Ala Asn Gly Thr Phe Asp Ala Met Lys Ser Met
140 145 150

Lys Thr Val His Leu Ala Lys Asn Pro Phe Ile Cys Asp Cys Asn
155 160 165

Leu Arg Trp Leu Ala Asp Tyr Leu His Lys Asn Pro Ile Glu Thr
170 175 180

Ser Gly Ala Arg Cys Glu Ser Pro Lys Arg Met His Arg Arg Arg
185 190 195

Ile Glu Ser Leu Arg Glu Glu Lys Phe Lys Cys Ser Trp Gly Glu
200 205 210

Leu Arg Met Lys Leu Ser Gly Glu Cys Arg Met Asp Ser Asp
215 220

(2) INFORMATION FOR SEQ ID NO: 5:
(i) SEQUENCE CHARACTERESITICS:
(A) LENGTH: 196 amino acids
(B) TYPE: amino acids
(D) TOPOLOGY: Linear
(ii) MOLECULE TYPE: protein
(ix) FEATURE:
(A) NAME/KEY: Flank-LRR-Flank 3
(B) LOCATION: 1 to 196
(C) IDENTIFICATION METHOD: similarity to other Flank-LRR-Flank domains defined herein.
(D) OTHER INFORMATION: mediates adhesive events
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

Cys Pro Ala Met Cys His Cys Glu Gly Thr Thr Val Asp Cys Thr
5 10 15

Gly Arg Arg Leu Lys Glu Ile Pro Arg Asp Ile Pro Leu His Thr
20 25 30

Thr Glu Leu Leu Leu Asn Asp Asn Glu Leu Gly Arg Ile Ser Ser
35 40 45

Asp Gly Leu Phe Gly Arg Leu Pro His Leu Val Lys Leu Glu Leu
50 55 60

Lys Arg Asn Gln Leu Thr Gly Ile Glu Pro Asn Ala Phe Glu Gly
65 70 75

Ala Ser His Ile Gln Glu Leu Gln Leu Gly Glu Asn Lys Ile Lys
80 85 90

Glu Ile Ser Asn Lys Met Phe Leu Gly Leu His Gln Leu Lys Thr
95 100 105

Leu Asn Leu Tyr Asp Asn Gln Ile Ser Cys Val Met Pro Gly Ser
110 115 120

Phe Glu His Leu Asn Ser Leu Thr Ser Leu Asn Leu Ala Ser Asn
125 130 135

Pro Phe Asn Cys Asn Cys His Leu Ala Trp Phe Ala Glu Cys Val
140 145 150

Arg Lys Lys Ser Leu Asn Gly Gly Ala Ala Arg Cys Gly Ala Pro
155 160 165

Ser Lys Val Arg Asp Val Gln Ile Lys Asp Leu Pro His Ser Glu
170 175 180

Phe Lys Cys Ser Ser Glu Asn Ser Glu Gly Cys Leu Gly Asp Gly
185 190 195
Tyr

(2) INFORMATION FOR SEQ ID NO: 6:
(i) SEQUENCE CHARACTERESITICS:
(A) LENGTH: 196 amino acids
(B) TYPE: amino acids
(D) TOPOLOGY: Linear
(ii) MOLECULE TYPE: protein
(ix) FEATURE:
(A) NAME/KEY: Flank-LRR-Flank 4
(B) LOCATION: 1 to 196
(C) IDENTIFICATION METHOD: similarity to other Flank-LRR-Flank domains defined herein.
(D) OTHER INFORMATION: mediates adhesive events

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Cys Pro Pro Ser Cys Thr Cys Thr Gly Thr Val Val Ala Cys Ser
5 10 15

Arg Asn Gln Leu Lys Glu Ile Pro Arg Gly Ile Pro Ala Glu Thr
20 25 30

Ser Glu Leu Tyr Leu Glu Ser Asn Glu Ile Glu Gln Ile His Tyr
35 40 45

Glu Arg Ile Arg His Leu Arg Ser Leu Thr Arg Leu Asp Leu Ser
50 55 60

Asn Asn Gln Ile Thr Ile Leu Ser Asn Tyr Thr Phe Ala Asn Leu
65 70 75

Thr Lys Leu Ser Thr Leu Ile Ile Ser Tyr Asn Lys Leu Gln Cys
80 85 90

Leu Gln Arg His Ala Leu Ser Gly Leu Asn Asn Leu Arg Val Val
95 100 105

Ser Leu His Gly Asn Arg Ile Ser Met Leu Pro Glu Gly Ser Phe
110 115 120

Glu Asp Leu Lys Ser Leu Thr His Ile Ala Leu Gly Ser Asn Pro
125 130 135

Leu Tyr Cys Asp Cys Gly Leu Lys Trp Phe Ser Asp Trp Ile Lys
140 145 150

Leu Asp Tyr Val Glu Pro Gly Ile Ala Arg Cys Ala Glu Pro Glu
155 160 165

Gln Met Lys Asp Lys Leu Ile Leu Ser Thr Pro Ser Ser Ser Phe
170 175 180

Val Cys Arg Gly Arg Val Arg Asn Asp Ile Leu Ala Lys Cys Asn
185 190 195

Ala

(2) INFORMATION FOR SEQ ID NO: 7:
(i) SEQUENCE CHARACTERESITICS:
(A) LENGTH: 11 amino acids
(B) TYPE: amino acids
(D) TOPOLOGY: Linear
(ii) MOLECULE TYPE: peptide
(v) FRAGMENT TYPE: Internal fragment
(ix) FEATURE:
(A) NAME/KEY: Alternate segment
(B) LOCATION: 1 to 11
(C) IDENTIFICATION METHOD: Experimental

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

Gly Glu Gly Ser Thr Glu Pro Pro Thr Val Thr
5 10

(2) INFORMATION FOR SEQ ID NO: 8:
(i) SEQUENCE CHARACTERESITICS:
(A) LENGTH: 196 amino acids
(B) TYPE: amino acids
(D) TOPOLOGY: Linear
(ii) MOLECULE TYPE: protein
(iii) HYPOTHETICAL: Yes
(ix) FEATURE:
(A) NAME/KEY: Flank-LRR-Flank consensus
(B) LOCATION: 1 to 196
(C) IDENTIFICATION METHOD: Experimental

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Cys Pro Xaa Xaa Cys Xaa Cys Xaa Gly Xaa Xaa Val Asp Cys Xaa
5 10 15

Xaa Xaa Gln Leu Xaa Xaa Xaa Pro Xaa Xaa Xaa Pro Xaa Asp Thr
20 25 30

Thr Xaa Xaa Xaa Xaa Xaa Xaa Asn Xaa Ile Xaa Xaa Leu Xaa Xaa
35 40 45

Xaa Xaa Phe Xaa Xaa Leu Xaa Xaa Leu Xaa Xaa Leu Xaa Leu Xaa
50 55 60

Xaa Asn Xaa Ile Xaa Xaa Leu Xaa Xaa Xaa Xaa Phe Xaa Xaa Leu
65 70 75

Xaa Xaa Leu Xaa Xaa Leu Ile Leu Xaa Xaa Asn Xaa Ile Xaa Xaa
80 85 90

Leu Xaa Xaa Xaa Xaa Leu Xaa Xaa Leu Xaa Xaa Leu Xaa Xaa Leu
95 100 105

Xaa Leu Xaa Xaa Asn Xaa Ile Xaa Xaa Leu Xaa Xaa Xaa Xaa Phe
110 115 120

Xaa Xaa Leu Xaa Xaa Leu Xaa Xaa Leu Xaa Leu Xaa Xaa Asn Pro
125 130 135

Phe Xaa Cys Asp Cys Xaa Leu Xaa Trp Leu Xaa Xaa Xaa Xaa Xaa
140 145 150

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Arg Cys Xaa Xaa Pro Xaa
155 160 165

Xaa Xaa Xaa Xaa Xaa Xaa Ile Xaa Xaa Leu Xaa Xaa Xaa Xaa Phe
170 175 180

Lys Cys Ser Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
185 190 195

Xaa

(2) INFORMATION FOR SEQ ID NO: 9:
(i) SEQUENCE CHARACTERESITICS:
(A) LENGTH: 76 amino acids
(B) TYPE: amino acids
(D) TOPOLOGY: Linear
(ii) MOLECULE TYPE: protein
(ix) FEATURE:
(A) NAME/KEY: COOH terminal region
(B) LOCATION: 1 to 76
(C) IDENTIFICATION METHOD: Experimental

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

Ala Ala Ser Thr Cys Arg Lys Glu Gln Val Arg Glu Tyr Tyr Thr
5 10 15

Glu Asn Asp Cys Arg Ser Arg Gln Pro Leu Lys Tyr Ala Lys Cys
20 25 30

Val Gly Gly Cys Gly Asn Gln Cys Cys Ala Ala Lys Ile Val Arg
35 40 45

Arg Arg Lys Val Arg Met Val Cys Ser Asn Asn Arg Lys Tyr Ile
50 55 60

Lys Asn Leu Asp Ile Val Arg Lys Cys Gly Cys Thr Lys Lys Cys
65 70 75

Tyr

WHAT IS CLAIMED IS:

1. An isolated and substantially pure form of the SLIT protein comprising the sequence SEQ I.D. NO. 2.

2. An isolated DNA segment encoding the entire SLIT protein, SEQ. I.D. NO. 1.

3. A recombinant expression vector comprising the DNA segment according to claim 2.

4. A recombinant host microorganism containing a DNA expression vector comprising a DNA sequence consisting essentially of a DNA sequence encoding the entire SLIT protein.

5. An amino flank-LRR-carboxy-flank concensus sequence element of the SLIT protein (SEQ. I.D. NO. 8) comprising

(a) an amino-flanking region comprising the sequence
CPxxCxC.....xGxxVDCxxxGLx...xαPxxαPxDTTx,

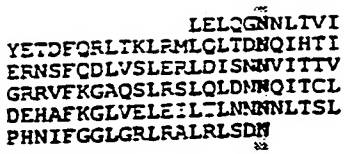
(b) a leucine-rich repeat region comprising one or more repeats of the sequence
xxxxFxxLxxLxLxxNxIxxL, and

(c) a carboxy-flanking region comprising the sequence
P(W or F)xC(D or N)Cxα.....W(L or F)xxxxxxxxxxxxxx..... RCxxPxxxxxxxαxxαxxxxFx..C (P or S).

6. The first amino-flank-LRR-carboxy-flank sequence element of the SLIT protein (SEQ. I.D. NO. 3) according to claim 5 wherein

(a) the amino-flanking region comprises the sequence
CPRVCSC TGLNVDCSHRGLT SVPRKISADVER,

(b) the leucine-rich region comprises the sequence



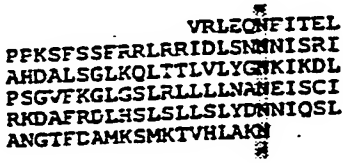
and

(c) the carboxy-flanking region comprises the sequence PFACD CHL SWLSRFLRSATRLAPYT RCQSPQLKGQNVADLHDQEFK CSGLTEHAPMECGAENS.

7. The second amino-flank-LRR-carboxy-flank sequence element of the SLIT protein (SEQ. I.D. NO. 4) according to claim 5 wherein

(a) the amino-flanking region comprises the sequence CPHPCRC ADGIVDCREKSLT SVPVTLPDDTTD,

(b) the leucine-rich region comprises the sequence



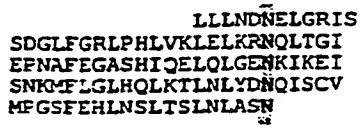
and

(c) the carboxy-flanking region comprises the sequence PFICNCNL RWLADYLHKIPIETSGARCESPKRMHRRIESLREEKFK CSWGELRMKLSGECRMDSD.

8. The third amino-flank-LRR-carboxy-flank sequence element of the SLIT protein (SEQ. I.D. NO. 5) according to claim 5, wherein

(a) the amino-flanking region comprises the sequence CPAMCHC EGTTVDCTGRGLK EIPRDIPLHTTE

(b) the leucine-rich repeat region comprises the sequence



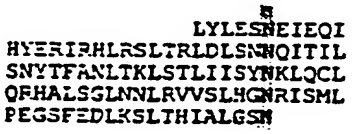
and

(c) the carboxy flanking region comprises the sequence PFNCNCHL AWFAECVRKKSLNGGAA RCGAPSKVRDVQIKDLPH SEEK CSSENSEGCLGD GY.

9. The fourth amino-flank-LRR-carboxy-flank sequence element of the SLIT protein (SEQ. I.D. NO. 6) according to claim 5, wherein

(a) the amino-flanking region comprising the sequence CPPSCTC TGTVVACSRNQLK EIPRGIPAETSE,

(b) the leucine-rich repeat region comprising the sequence



and

(c) the carboxy-flanking region comprising the sequence PLYCDCGL KWFSDWIKLDYVEPGIA RCAEPEQMKDKLILSTPSSSFV CRGRVRNDILAKCNA.

10. The alternate splice segment of the SLIT protein residing at the seventh epidermal growth factor (EGF) sequence element of the SLIT protein comprising the sequence GEGSTEPFTVT (SEQ. I.D. NO. 7).

11. The carboxy terminal region of the SLIT protein (SEQ. I.D. NO. 9) residing after the seventh epidermal growth factor.

12. A combination comprising one or more amino-flank-LRR-carboxy-flank sequence elements according to claim 5 and one or more EGF-like repeat elements of the SLIT protein, provided that said combination does not include the naturally occurring configuration of the SLIT protein.

13. The combination according to claim 12, further including the alternative splice segment of the SLIT protein residing at the seventh epidermal growth factor sequence element when part of the SLIT protein comprising the sequence GEGSTEPFTVT (SEQ. I.D. NO. 7).

14. An antibody to the SLIT protein produced by introducing into an animal the SLIT protein according to claim 1 sufficient to raise antibodies in said animal and withdrawing said antibodies from said animal.

15. An antibody to the SLIT protein produced by introducing into an animal the SLIT protein according to claim 5 sufficient to raise antibodies in said animal and withdrawing said antibodies from said animal.

16. An antibody to the SLIT protein produced by introducing into an animal the SLIT protein according to claim 6 sufficient to raise antibodies in said animal and withdrawing said antibodies from said animal.

17. An antibody to the SLIT protein produced by introducing into an animal the SLIT protein according to claim 7 sufficient to raise antibodies in said animal and withdrawing said antibodies from said animal.

18. An antibody to the SLIT protein produced by introducing into an animal the SLIT protein according to claim 8 sufficient to raise antibodies in said animal and withdrawing said antibodies from said animal.

19. An antibody to the SLIT protein produced by introducing into an animal the SLIT protein according to claim 9 sufficient to raise antibodies in said animal and withdrawing said antibodies from said animal.

20. An antibody to the SLIT protein produced by introducing into an animal the SLIT protein according to claim 10 sufficient to raise antibodies in said animal and withdrawing said antibodies from said animal.

21. An antibody to the SLIT protein produced by introducing into an animal the SLIT protein according to claim 11 sufficient to raise antibodies in said animal and withdrawing said antibodies from said animal.

22. An antibody to the SLIT protein produced by introducing into an animal the SLIT protein according to claim 12 sufficient to raise antibodies in said animal and withdrawing said antibodies from said animal.

23. A method of detecting the SLIT protein or a shed portion thereof in a bodily fluid comprising contacting the bodily fluid with antibodies raised to the SLIT protein according to claim 1 or to a portion thereof and detecting for the presence of the SLIT protein.

24. A method of detecting autoimmune antibodies to the SLIT protein or a shed portion thereof in a bodily fluid comprising contacting the bodily fluid with the SLIT protein according to claim 1 or a portion thereof and detecting for the presence of said autoimmune antibodies.

25. A method of detecting chromosomal rearrangements in the SLIT locus comprising hybridizing a nucleic acid from a patient with a nucleic acid sequence from the SLIT locus and detecting for the level of expression or an aberrant rearrangement, said nucleic acid sequence being the DNA according to claim 2 or a portion thereof.

26. A pharmaceutical preparation for the treatment of a neurodegenerative disease, for treating a traumatic injury to a neural tissue or for affecting the angiogenic process in a patient comprising a sterile preparation comprising an effective amount of the SLIT protein according to claim 1 or a portion thereof, in admixture with a pharmaceutically acceptable carrier.

27. A pharmaceutical preparation for the treatment of a neurodegenerative disease, for treating a traumatic injury to a neural tissue or for affecting the angiogenic process in a patient comprising a sterile preparation comprising an effective amount of the SLIT protein according to claim 5, in admixture with a pharmaceutically acceptable carrier.

28. A pharmaceutical preparation for the treatment of a neurodegenerative disease, for treating a traumatic injury to a neural tissue or for affecting the angiogenic process in a patient comprising a sterile preparation comprising an effective amount of the SLIT protein according to claim 6, in admixture with a pharmaceutically acceptable carrier.

29. A pharmaceutical preparation for the treatment of a neurodegenerative disease, for treating a traumatic injury to a neural tissue or for affecting the angiogenic process in a patient comprising a sterile preparation comprising an effective amount of the SLIT protein according to claim 7, in admixture with a pharmaceutically acceptable carrier.

30. A pharmaceutical preparation for the treatment of a neurodegenerative disease, for treating a traumatic injury to a neural tissue or for affecting the angiogenic process in a patient comprising a sterile preparation comprising an effective amount of the SLIT protein according to claim 8, in admixture with a pharmaceutically acceptable carrier.

31. A pharmaceutical preparation for the treatment of a neurodegenerative disease, for treating a traumatic injury to a neural tissue or for affecting the angiogenic process in a patient comprising a sterile preparation comprising an effective amount of the SLIT protein according to claim 9, in admixture with a pharmaceutically acceptable carrier.

32. A pharmaceutical preparation for the treatment of a neurodegenerative disease, for treating a traumatic injury to a neural tissue or for affecting the angiogenic process in a patient comprising a sterile preparation comprising an effective amount of the SLIT protein according to claim 10, in admixture with a pharmaceutically acceptable carrier.

33. A pharmaceutical preparation for the treatment of a neurodegenerative disease, for treating a traumatic injury to a neural tissue or for affecting the angiogenic process in a patient comprising a sterile preparation comprising an effective amount of the SLIT protein according to claim 11, in admixture with a pharmaceutically acceptable carrier.

34. A pharmaceutical preparation for the treatment of a neurodegenerative disease, for treating a traumatic injury to a neural tissue or for affecting the angiogenic process in a patient comprising a sterile preparation comprising an effective amount of the SLIT protein according to claim 12, in admixture with a pharmaceutically acceptable carrier.

35. A method for the treatment of a neurodegenerative disease, for treating a traumatic injury to a neural tissue or for affecting the angiogenic process in a patient comprising administering to said patient an effective amount of the SLIT protein according to claim 1, or a portion thereof, either alone or in admixture with a pharmaceutically acceptable carrier.

36. A method for the treatment of a neurodegenerative disease, for treating a traumatic injury to a neural tissue or for affecting the angiogenic process in a patient comprising administering to said patient an effective amount of the SLIT protein according to claim 5, either alone or in admixture with a pharmaceutically acceptable carrier.

37. A method for the treatment of a neurodegenerative disease, for treating a traumatic injury to a neural tissue or for affecting the angiogenic process in a patient comprising administering to said patient an effective amount of the SLIT protein according to claim 6, either alone or in admixture with a pharmaceutically acceptable carrier.

38. A method for the treatment of a neurodegenerative disease, for treating a traumatic injury to a neural tissue or for affecting the angiogenic process in a patient comprising administering to said patient an effective amount of the SLIT protein according to claim 7, either alone or in admixture with a pharmaceutically acceptable carrier.

39. A method for the treatment of a neurodegenerative disease, for treating a traumatic injury to a neural tissue or for affecting the angiogenic process in a patient comprising administering to said patient an effective amount of the SLIT protein according to claim 8, either alone or in admixture with a pharmaceutically acceptable carrier.

40. A method for the treatment of a neurodegenerative disease, for treating a traumatic injury to a neural tissue or for affecting the angiogenic process in a patient comprising administering to said patient an effective amount of the SLIT protein according to claim 9, either alone or in admixture with a pharmaceutically acceptable carrier.

41. A method for the treatment of a neurodegenerative disease, for treating a traumatic injury to a neural tissue or for affecting the angiogenic process in a patient comprising administering to said patient an effective amount of the SLIT protein according to claim 10, either alone or in admixture with a pharmaceutically acceptable carrier.

42. A method for the treatment of a neurodegenerative disease, for treating a traumatic injury to a neural tissue or for affecting the angiogenic process in a patient comprising administering to said patient an effective amount of the SLIT protein according to claim 11, either alone or in admixture with a pharmaceutically acceptable carrier.

43. A method for the treatment of a neurodegenerative disease, for treating a traumatic injury to a neural tissue or for affecting the angiogenic process in a patient comprising administering to said patient an effective amount of the SLIT protein according to claim 12, either alone or in admixture with a pharmaceutically acceptable carrier.

44. A protein, TAGON, that allows for the formation of a molecular bridge between axonally associated receptors and extracellular matrix molecules.

45. A pharmaceutical preparation for the treatment of a neurodegenerative disease, for treating a traumatic injury to a neural tissue or for affecting the angiogenic process in a patient comprising a sterile preparation comprising an effective amount of a TAGON protein in admixture with a pharmaceutically acceptable carrier.

46. A method for the treatment of a neurodegenerative disease, for treating a traumatic injury to a neural tissue or

for affecting the angiogenic process in a patient comprising administering to said patient an effective amount of a TAGON protein, either alone or in admixture with a pharmaceutically acceptable carrier.

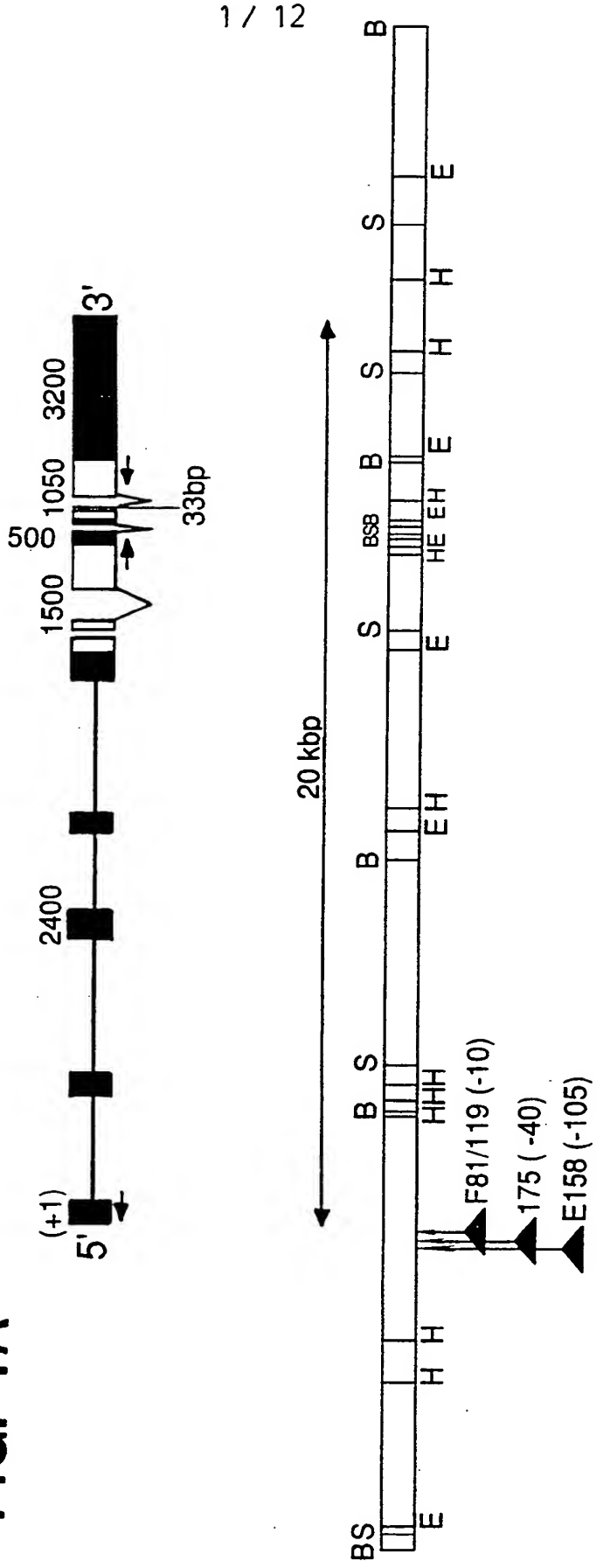

3'
3200
1050
33bp
500
1500
2400
(+1)
5'
20 kbp
B
S
E
H
BSB
HE EH
HHH
F81/119 (-10)
175 (-40)
E158 (-105)
BS

FIG. 1A

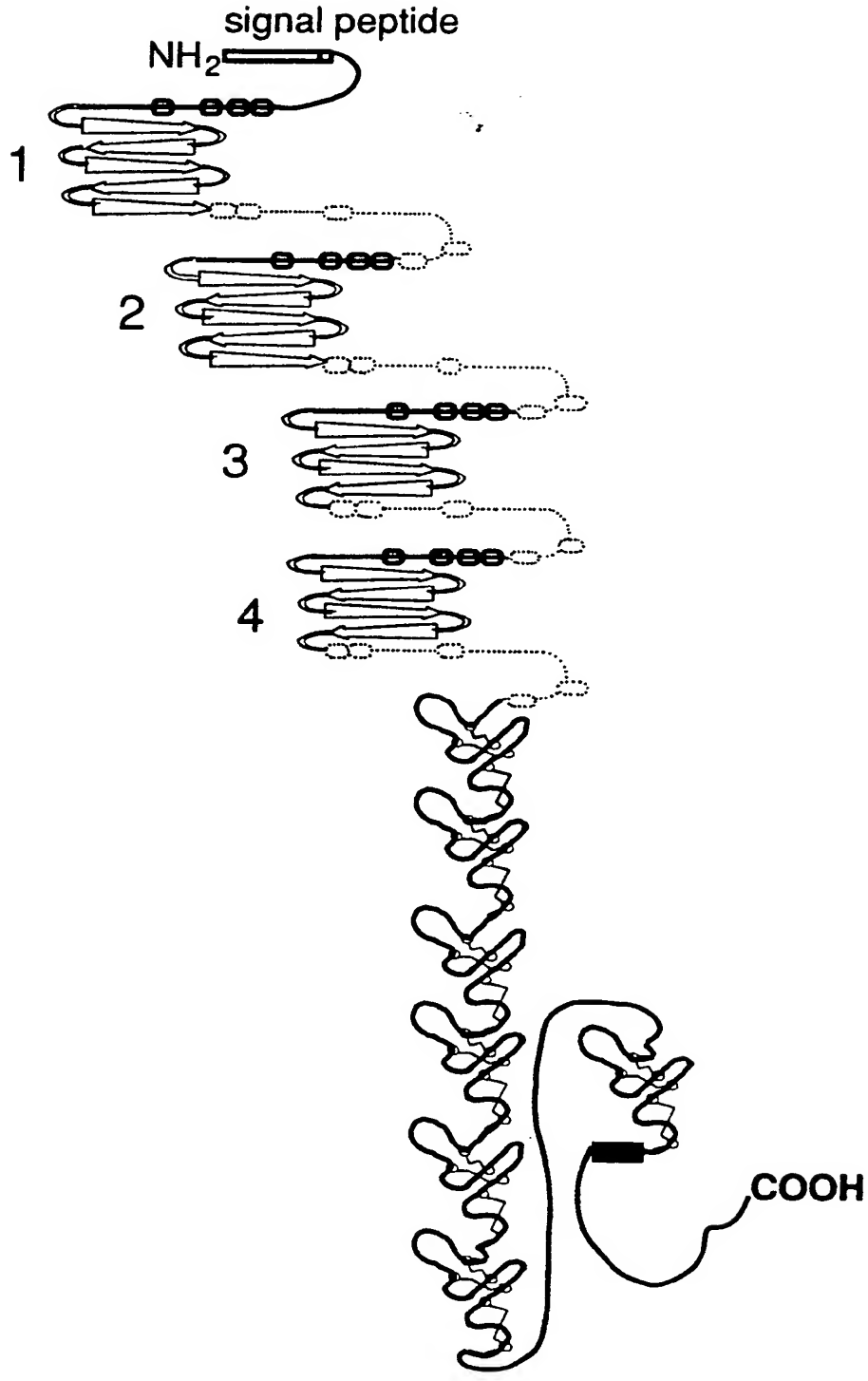


FIG. 2A

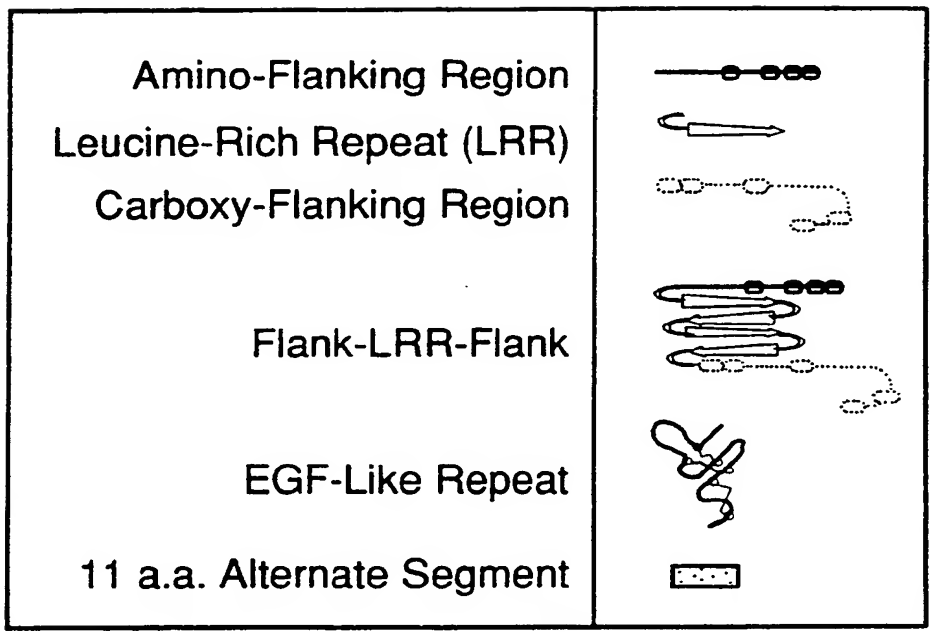

Amino-Flanking Region
Leucine-Rich Repeat (LRR)
Carboxy-Flanking Region
Flank-LRR-Flank
EGF-Like Repeat
11 a.a. Alternate Segment

FIG. 2B

FIG. 3C

FIG. 3F

FIG. 3B

FIG. 3E

FIG. 3A

FIG. 3D

FIG. 3I

FIG. 3L

FIG. 3H

FIG. 3K

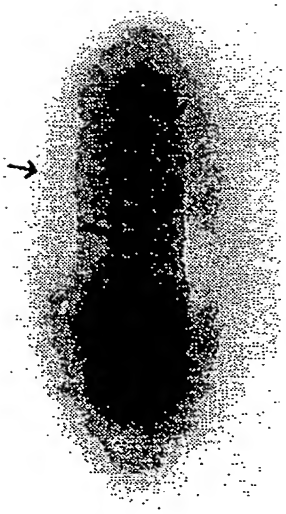


FIG. 3G

FIG. 3J

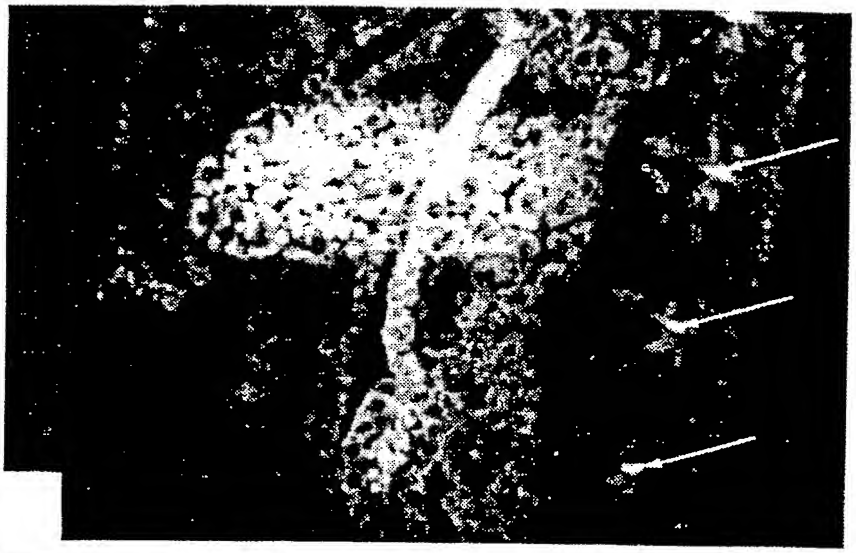


FIG. 4C

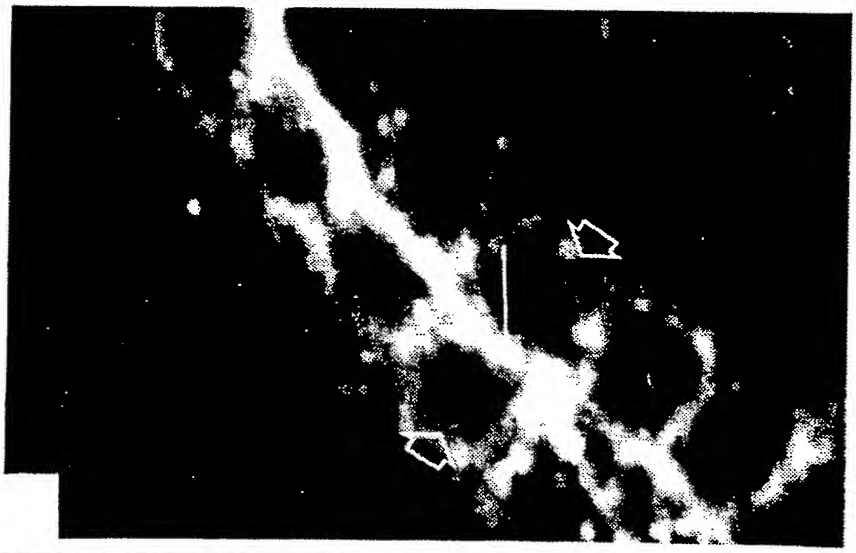


FIG. 4B

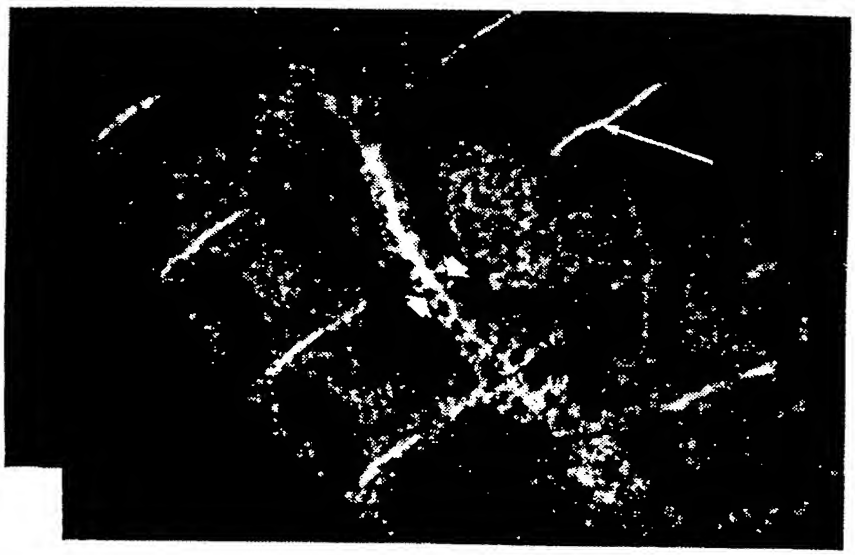


FIG. 4A

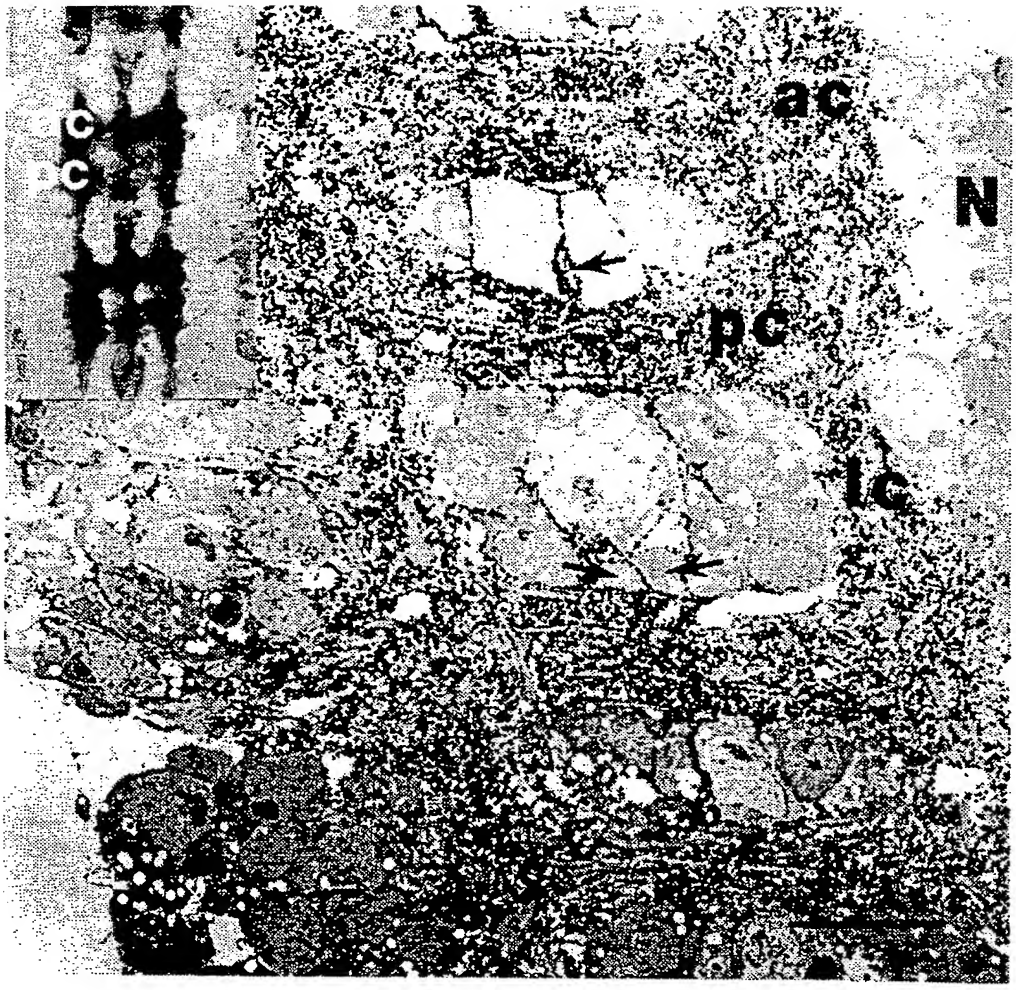


FIG. 5

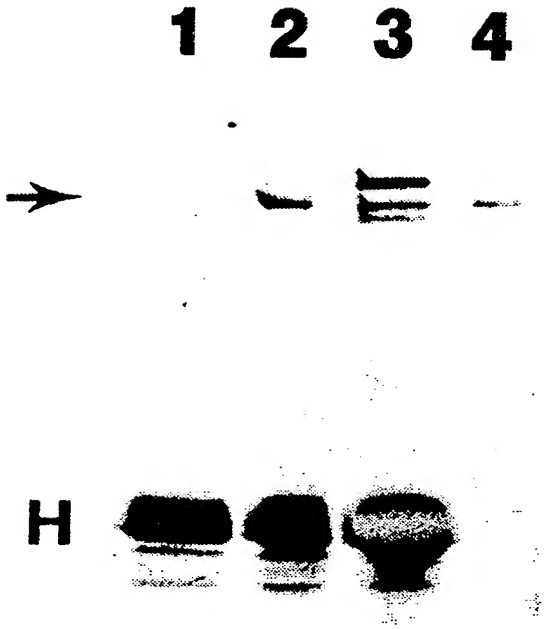


FIG. 6A

FIG. 6B

FIG. 7B

FIG. 7D

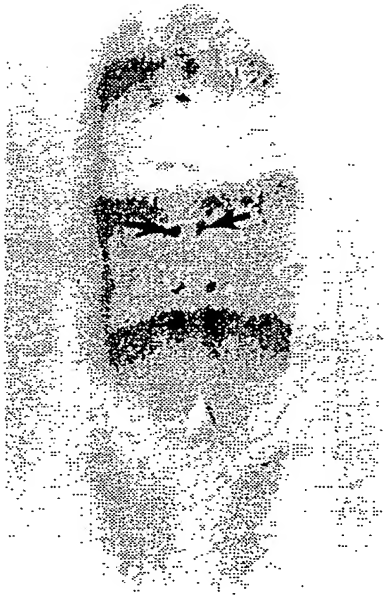


FIG. 7A

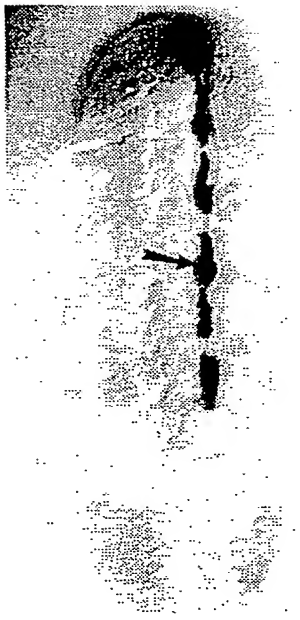


FIG. 7C

FIG. 7F

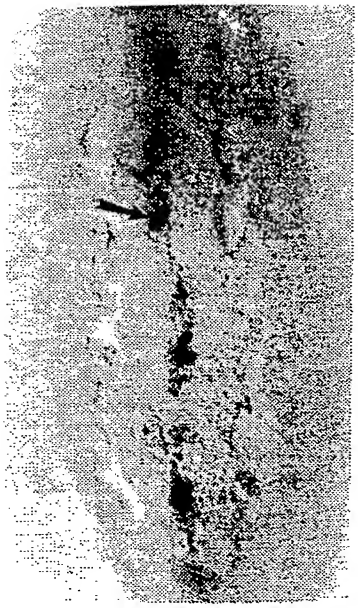


FIG. 7H

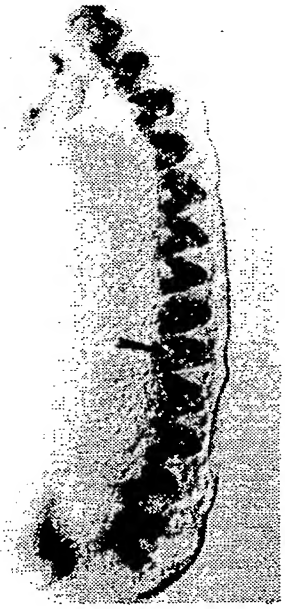


FIG. 7E

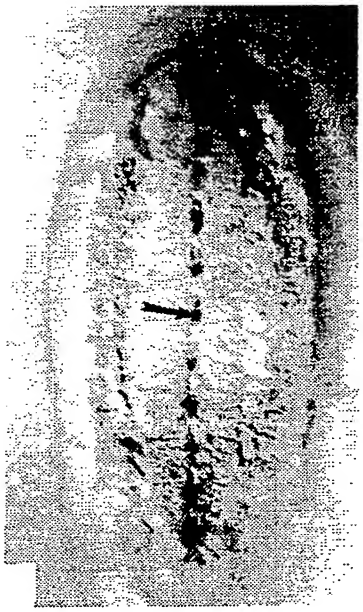


FIG. 7G

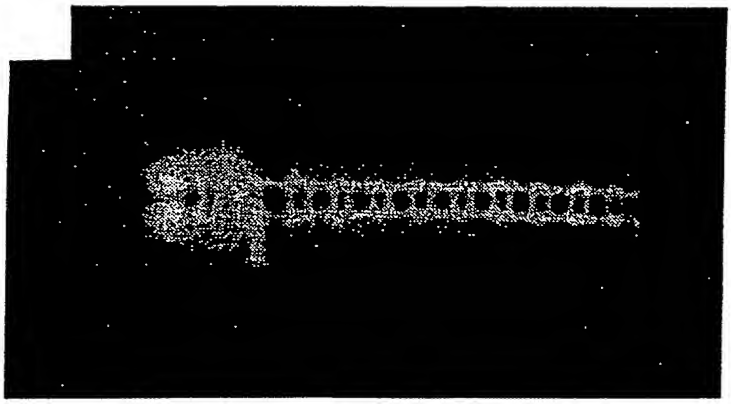


FIG. 8A

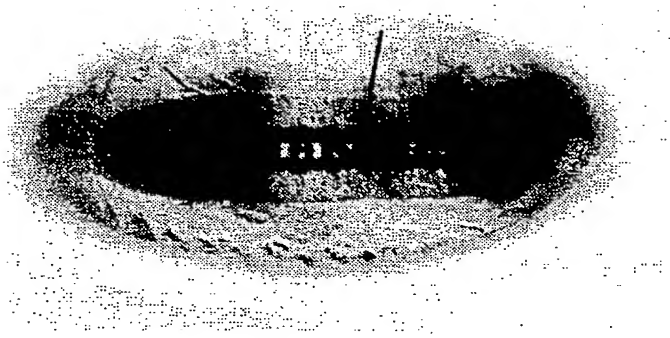


FIG. 8B

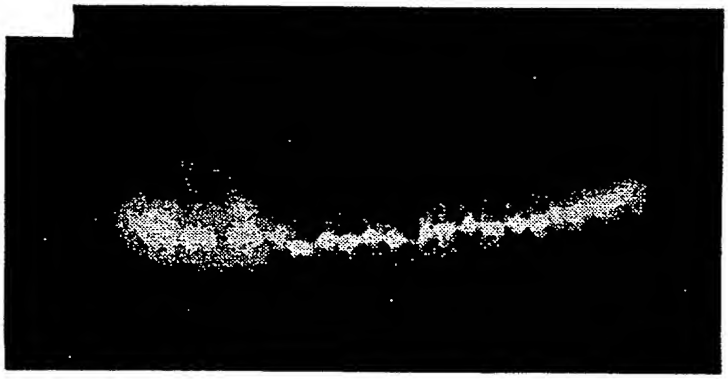


FIG. 8C

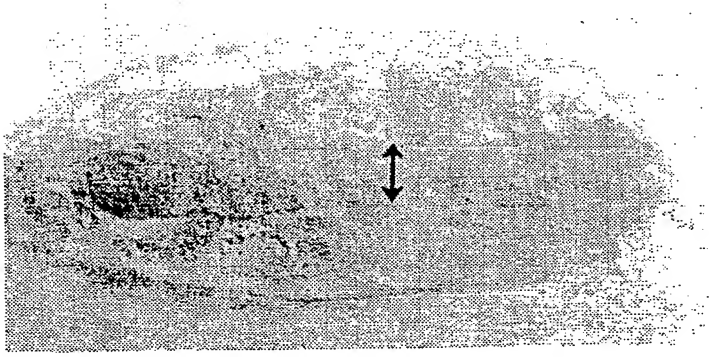


FIG. 8D

FIG. 8E

FIG. 8F

INTERNATIONAL SEARCH REPORT

International Application No. PCT/US91/09055

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all)[3]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC (5): C07K 13/00
US CL : 530/350

II. FIELDS SEARCHED

Minimum Documentation Searched[4]

Classification System	Classification Symbols
U.S.	530/350

Documentation Searched other than Minimum Documentation to the extent that such Documents are included in the Fields Searched[5]

APS, Medline
search terms: slit protein, Rothberg JM, neuron?(p)adhesion

III. DOCUMENTS CONSIDERED TO BE RELEVANT[14]

Category*	Citation of Document,[16] with indication, where appropriate, of the relevant passages[17]	Relevant to Claim No.[18]
Y	Cell, Volume 55, issued 22 December 1988, J.M. Rothberg et al., "slit: An EGF-homologous locus of D. melanogaster involved in the development of the embryonic central nervous system," pages 1047-1059, see entire contents.	1

* Special categories of cited documents:[15]

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search[2]	Date of Mailing of this International Search Report[2]
26 MARCH 1992	08 APR 1992
International Searching Authority[1]	**Signature of Authorized Officer[20]**
ISA/US	Michael P. Woodward

Form PCT/ISA/210 (second sheet)(May 1986) ß

FURTHER INFORMATION CONTINUED FROM PREVIOUS SHEETS
(Not for publication)

VI. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

I. Claim 1, drawn to isolated slit protein, classified in Class 530, subclass 350.
II. Claims 2-4 drawn to an isolated DNA sequence, an expression vector and transformed host cell, classified in Class 536, subclass 27 and 435, subclass 320.1 respectively.
III. Claims 5-13, drawn to peptides, classified in Class 530, subclass 350 and Class 530, subclass 324 or 327.
IV. Claims 14-22, drawn to antibodies, classified in Class 530, subclass 387.
V. Claims 26-34, drawn to pharmaceutical compositions, classified in Class 514, subclass 12.
VI. Claims 35-43, drawn to treatment methods involving peptides, classified in Class 514, subclas 2.
VII. Claim 23, drawn to an immunoassay for slit protein, classified in Class 435, subclass 7.1.
VIII. Claim 24, drawn to an immun oassay for anti-slit protein antibodies, classified in Class 435, subclass 7.1.
IX. Claim 25, drawn to a hybridization assay, classified in Class 435, sublcass 6.
X. Claim 44, drawn to tagon protein, classified in Class 530, subclass 350.
XI. Claim 45, drawn to a pharmaceutical composition containing tagon, classified in Class 514, subclass 2.
XII. Claim 46, drawn to a mthod of treatment using tagon protein, classified in Class 514, subclass 2.

The claims of these twelve groups are drawn to distinct inventions which are not linked so as to form a single general inventive concept. PCT rules 13.1 and 13.2 do not provide for multiple products and methods. The claims of groups III-VI have been grouped using PCT Rule 13.3.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :		(11) International Publication Number:	WO 92/10518
C07K 13/00	A1	(43) International Publication Date:	25 June 1992 (25.06.92)

(21) International Application Number: PCT/US91/09055

(22) International Filing Date: 27 November 1991 (27.11.91)

(30) Priority data:
624,135 7 December 1990 (07.12.90) US

(71) Applicant: YALE UNIVERSITY [US/US]; 246 Church Street, New Haven, CT 06510 (US).

(71)(72) Applicants and Inventors: ROTHBERG, Jonathan, Marc [US/US]; 642 Rosemount Lane, Westville, CT 06515 (US). ARTAVANIS-TSAKONAS, Spyridon [US/US]; 192 Ridgewood Avenue, Hamden, CT 06517 (US).

(74) Agent: BARTH, Richard, S.; Frishauf, Holtz, Goodman & Woodward, 600 Third Avenue, New York, NY 10016 (US).

(81) Designated States: AT (European patent), BE (European patent), CA, CH (European patent), DE (European patent), DK (European patent), ES (European patent), FR (European patent), GB (European patent), GR (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent).

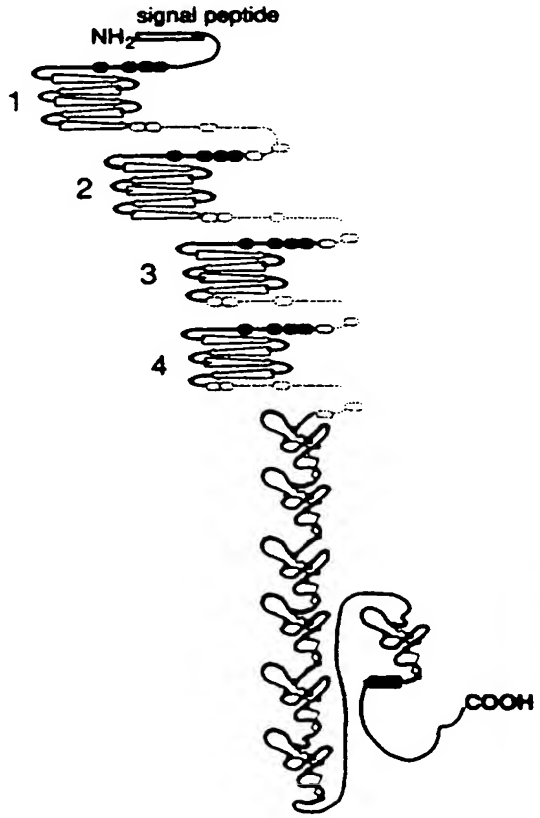
Published
With international search report.
With amended claims.

Date of publication of the amended claims:
20 August 1992 (20.08.92)

(54) Title: PURIFIED SLIT PROTEIN AND SEQUENCE ELEMENTS THEREOF

(57) Abstract

An isolated and substantially pure form of the SLIT protein and sequence elements thereof, antibodies thereto and diagnostics and therapeutics utilizing such proteins and antibodies. A method for treating neurodegenerative disease, traumatic injury to a neural tissue or affecting the angiogenic process in a patient comprising administering to the patient an effective amount of the SLIT protein.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**FI**	Finland	**ML**	Mali
AU	Australia	**FR**	France	**MN**	Mongolia
BB	Barbados	**GA**	Gabon	**MR**	Mauritania
BE	Belgium	**GB**	United Kingdom	**MW**	Malawi
BF	Burkina Faso	**GN**	Guinea	**NL**	Netherlands
BG	Bulgaria	**GR**	Greece	**NO**	Norway
BJ	Benin	**HU**	Hungary	**PL**	Poland
BR	Brazil	**IE**	Ireland	**RO**	Romania
CA	Canada	**IT**	Italy	**RU**	Russian Federation
CF	Central African Republic	**JP**	Japan	**SD**	Sudan
CG	Congo	**KP**	Democratic People's Republic of Korea	**SE**	Sweden
CH	Switzerland			**SN**	Senegal
CI	Côte d'Ivoire	**KR**	Republic of Korea	**SU**	Soviet Union
CM	Cameroon	**LI**	Liechtenstein	**TD**	Chad
CS	Czechoslovakia	**LK**	Sri Lanka	**TG**	Togo
DE	Germany	**LU**	Luxembourg	**US**	United States of America
DK	Denmark	**MC**	Monaco		
ES	Spain	**MG**	Madagascar		

AMENDED CLAIMS

[received by the International Bureau on 4 June 1992 (04.06.92); original claims 1-46 replaced by amended claims 1-93 (13 pages)]

1. An isolated and substantially pure SLIT protein comprising the amino acid sequence as depicted in SEQ ID NO:2 from amino acid numbers 1 through 1480.

2. An isolated and substantially pure SLIT protein comprising the amino acid sequence as depicted in SEQ ID NO:2 from amino acid numbers 37 through 1480.

3. An isolated and substantially pure SLIT protein comprising the amino acid sequence as depicted in SEQ ID NO:2 from amino acid numbers 37 through 1393 fused to the amino acid sequence as depicted in SEQ ID NO:2 from amino acid numbers 1405 through 1480.

4. An isolated and substantially pure SLIT protein comprising the amino acid sequence as depicted in SEQ ID NO:2 from amino acid numbers 1 through 1393 fused to the amino acid sequence as depicted in SEQ ID NO:2 from amino acid numbers 1405 through 1480.

5. An isolated and substantially pure protein comprising the amino acid sequence as depicted in SEQ ID NO:2 from amino acid numbers 1394 through 1480.

6. An isolated and substantially pure protein comprising the amino acid sequence as depicted in SEQ ID NO:3 from amino acid numbers 1 through 222.

7. An isolated and substantially pure protein comprising the amino acid sequence as depicted in SEQ ID NO:4 from amino acid numbers 1 through 224.

8. An isolated and substantially pure protein comprising the amino acid sequence as depicted in SEQ ID NO:5 from amino acid numbers 1 through 196.

9. An isolated and substantially pure protein comprising the amino acid sequence as depicted in SEQ ID NO:6 from amino acid numbers 1 through 196.

10. An isolated and substantially pure protein comprising the amino acid sequence as depicted in SEQ ID NO:8 from amino acid numbers 1 through 196.

11. An isolated and substantially pure protein comprising the amino acid sequence as depicted in SEQ ID NO:9 from amino acid numbers 1 through 76.

12. An isolated and substantially pure protein comprising the amino acid sequence

P(W or F)XC(D or N)CXXW
(L or F)XXXXXXXXXXXXXXRCXXPX
XXXXXXXXXXXXXXFXC(P or S).

13. An isolated and substantially pure protein comprising:

(a) the amino acid sequence

CPXXCXCXGXXVDCX
XXGLXXXPXXXPXDTTX; and

(b) the amino acid sequence

P(W or F)XC(D or N)CXXW
(L or F)XXXXXXXXXXXXXXRCXXPX
XXXXXXXXXXXXXXFXC(P or S).

14. An isolated and substantially pure protein comprising:

(a) the amino acid sequence

XXXXFXXLXXLXXLXLXXNXIXXL; and

(b) the amino acid sequence

P(W or F)XC(D or N)CXXW
(L or F)XXXXXXXXXXXXXXRCXXPX
XXXXXXXXXXXXXXFXC(P or S).

15. An isolated and substantially pure protein comprising an amino acid sequence selected from the group consisting of:

CPRVCSC TGLNVDCSHRGLT SVPRKISADVER;

CPHPCRC ADGIVDCREKSLT SVPVTLPDDTTD;

CPAMCHC EGTTVDCTGRRLK EIPRDIPLHTTE; and

CPPSCTC TGTVVACSRNQLK EIPRGIPAETSE.

16. An isolated and substantially pure protein comprising an amino acid sequence selected from the group consisting of:

LELQGNNLTVI
YETDFQRLTKLRMLQLTDNQIHTI
ERNSFQDLVSLERLDISNNVITTV
GRRVFKGAQSLRSLQLDNNQITCL
DEHAFKGLVELEILTLNNNNLTSL
PHNIFGGLGRLRALRLSDN;

VRLEQNFITEL
PPKSFSSFRRLRRIDLSNNNISRI
AHDALSGLKQLTTLVLYGNKIKDL
PSGVFKGLGSLRLLLLNANEISCI
RKDAFRDLHSLSLLSLYDNNIQSL
ANGTFDAMKSMKTVHLAKN;

LLLNDNELGRIS
SDGLFGRLPHLVKLELKRNQLTGI
EPNAFEGASHIQELQLGENKIKEI
SNKMFLGLHQLKTLNLYDNQISCV
MPGSFEHLNSLTSLNLASN; and

LYLESNEIEQI
HYERIRHLRSLTRLDLSNNQITIL
SNYTFANLTKLSTLIISYNKLQCL
QRHALSGLNNLRVVSLHGNRISML
PEGSFEDLKSLTHIALGSN.

17. An isolated and substantially pure protein comprising an amino acid sequence selected from the group consisting of

PFACDCHLSWLSRFLRSAT
RLAPYTRCQSPSQLKGQNVADLHD
QEFKCSGLTEHAPMECGAENS;

PFICDCNLRWLADYLHKNP
IETSGARCESPKRMHRRRIESLRE
EKFKCSWGELRMKLSGECRMDSD;

PFNCNCHLAWFAECVRKKS
LNGGAARCGAPSKVRDVQIKDLPH
SEFKCSSENSEGCLGDGY; and

PLYCDCGLKWFSDWIKLDY
VEPGIARCAEPEQMKDKLILSTPS
SSFVCRGRVRNDILAKCNA.

18. The protein of claim 10 which further comprises an epidermal growth factor-like repeat, with the proviso that said protein is not the SLIT protein of claim 1, 2, 3 or 4.

19. The protein of claim 12 which further comprises an epidermal growth factor-like repeat, with the proviso that said protein is not the SLIT protein of claim 1, 2, 3 or 4.

20. The protein of claim 13 which further comprises an epidermal growth factor-like repeat, with the proviso that said protein is not the SLIT protein of claim 1, 2, 3 or 4.

21. The protein of claim 14 which further comprises an epidermal growth factor-like repeat, with the proviso that said protein is not the SLIT protein of claim 1, 2, 3 or 4.

22. The protein of claim 16 which further comprises an epidermal growth factor-like repeat, with the proviso that said protein is not the SLIT protein of claim 1, 2, 3 or 4.

23. The protein of claim 18 in which the epidermal growth factor-like repeat is selected from the group consisting of a 40 amino acid subsequence of the sequence as depicted in SEQ ID NO:2 from amino acid numbers 911 through 1150, and the amino acid sequence as depicted in SEQ ID NO:2 from amino acid numbers 1353 through 1393.

24. The protein of claim 19 in which the epidermal growth factor-like repeat is selected from the group consisting of a 40 amino acid subsequence of the sequence as

depicted in SEQ ID NO:2 from amino acid numbers 911 through 1150, and the amino acid sequence as depicted in SEQ ID NO:2 from amino acid numbers 1353 through 1393.

25. The protein of claim 20 in which the epidermal growth factor-like repeat is selected from the group consisting of a 40 amino acid subsequence of the sequence as depicted in SEQ ID NO:2 from amino acid numbers 911 through 1150, and the amino acid sequence as depicted in SEQ ID NO:2 from amino acid numbers 1353 through 1393.

26. The protein of claim 21 in which the epidermal growth factor-like repeat is selected from the group consisting of a 40 amino acid subsequence of the sequence as depicted in SEQ ID NO:2 from amino acid numbers 911 through 1150, and the amino acid sequence as depicted in SEQ ID NO:2 from amino acid numbers 1353 through 1393.

27. The protein of claim 22 in which the epidermal growth factor-like repeat is selected from the group consisting of a 40 amino acid subsequence of the sequence as depicted in SEQ ID NO:2 from amino acid numbers 911 through 1150, and the amino acid sequence as depicted in SEQ ID NO:2 from amino acid numbers 1353 through 1393.

28. An isolated and substantially pure protein comprising:

(a) the amino acid sequence

CPXXCXCXGXXVDCXXX
GLXXXPXXXPXDTTX;

(b) the amino acid sequence

XXXXFXXLXXLXXLXLXXNXIXXL; and

(c) the amino acid sequence

P(W or F)XC(D or N)CXXW
(L or F)XXXXXXXXXXXXXXRCXXPX
XXXXXXXXXXXXXXFXC(P or S).

29. An isolated and substantially pure protein having the amino acid sequence as depicted in SEQ ID NO:3 from amino acid numbers 1 through 222.

30. An isolated and substantially pure protein having the amino acid sequence as depicted in SEQ ID NO:4 from amino acid numbers 1 through 224.

31. An isolated and substantially pure protein having the amino acid sequence as depicted in SEQ ID NO:5 from amino acid numbers 1 through 196.

32. An isolated and substantially pure protein having the amino acid sequence as depicted in SEQ ID NO:6 from amino acid numbers 1 through 196.

33. A method for producing a SLIT protein comprising:
(a) growing a recombinant cell containing a nucleic acid vector encoding the SLIT protein of claim 1, whereby the SLIT protein is expressed by the cell; and
(b) obtaining the expressed SLIT protein.

34. A method for producing a SLIT protein comprising:
(a) growing a recombinant cell containing a nucleic acid vector encoding the SLIT protein of claim 2, whereby the SLIT protein is expressed by the cell; and
(b) obtaining the expressed SLIT protein.

35. A method for producing a SLIT protein comprising:
(a) growing a recombinant cell containing a nucleic acid vector encoding the SLIT protein of claim 3, whereby the SLIT protein is expressed by the cell; and
(b) obtaining the expressed SLIT protein.

36. A method for producing a SLIT protein comprising:
(a) growing a recombinant cell containing a nucleic acid vector encoding the SLIT protein of claim 4, whereby the SLIT protein is expressed by the cell; and
(b) obtaining the expressed SLIT protein.

37. The method according to claim 33 in which the SLIT protein is secreted by the cell.

38. The method according to claim 34 in which the SLIT protein is secreted by the cell.

39. The method according to claim 35 in which the SLIT protein is secreted by the cell.

40. The method according to claim 36 in which the SLIT protein is secreted by the cell.

41. The method according to claim 33 in which the SLIT protein is glycosylated by the cell.

42. The method according to claim 34 in which the SLIT protein is glycosylated by the cell.

43. An isolated and substantially pure SLIT protein produced according to the method of claim 37.

44. An isolated and substantially pure SLIT protein produced according to the method of claim 38.

45. An isolated and substantially pure SLIT protein produced according to the method of claim 39.

46. An isolated and substantially pure SLIT protein produced according to the method of claim 40.

47. An antibody which binds to a protein having the amino acid sequence as depicted in SEQ ID NO:2 from amino acid numbers 1 through 1150, or amino acid numbers 37 through 1150.

48. An antibody which binds to a protein having the amino acid sequence as depicted in SEQ ID NO:2 from amino acid numbers 1394 through 1480, or amino acid numbers 1405 through 1480.

49. An isolated DNA molecule comprising the DNA sequence as depicted in SEQ ID NO:1 from nucleotide numbers 315 through 4754.

50. An isolated nucleic acid comprising a nucleotide sequence encoding the protein of claim 1.

51. An isolated nucleic acid comprising a nucleotide sequence encoding the protein of claim 2.

52. An isolated nucleic acid comprising a nucleotide sequence encoding the protein of claim 3.

53. An isolated nucleic acid comprising a nucleotide sequence encoding the protein of claim 4.

54. An isolated nucleic acid comprising a nucleotide sequence encoding the protein of claim 5.

55. An isolated nucleic acid comprising a nucleotide sequence encoding an amino acid sequence as depicted in SEQ ID NO:3 from amino acid numbers 1 through 222, as depicted in SEQ ID NO:4 from amino acid numbers 1 through 224, as depicted in SEQ ID NO:5 from amino acid numbers 1 through 196, or as depicted in SEQ ID NO:6 from amino acid numbers 1 through 196.

56. An isolated nucleic acid comprising a nucleotide sequence encoding the protein of claim 10.

57. An isolated nucleic acid comprising a nucleotide sequence encoding the protein of claim 11.

58. An isolated nucleic acid comprising a nucleotide sequence encoding the protein of claim 12.

59. An isolated nucleic acid comprising a nucleotide sequence encoding the protein of claim 18.

60. An isolated nucleic acid comprising a nucleotide sequence encoding the protein of claim 23.

61. An isolated nucleic acid comprising a nucleotide sequence encoding the protein of claim 24.

62. The nucleic acid of claim 51 which is a cDNA molecule.

63. The nucleic acid of claim 52 which is a cDNA molecule.

64. The nucleic acid of claim 56 which is a cDNA molecule.

65. The nucleic acid of claim 57 which is a cDNA molecule.

66. The nucleic acid of claim 61 which is a cDNA molecule.

67. An isolated nucleic acid complementary to the cDNA molecule of claim 62.

68. An isolated nucleic acid complementary to the cDNA molecule of claim 66.

69. A recombinant expression vector comprising the DNA molecule of claim 49.

70. A recombinant host cell containing a nucleic acid vector comprising the nucleic acid of claim 50, 53 or 54.

71. A recombinant host cell containing a nucleic acid vector comprising the cDNA molecule of claim 62, 63 or 64.

72. A recombinant host cell containing a nucleic acid vector comprising the cDNA molecule of claim 65 or 66.

73. A recombinant host cell containing a nucleic acid vector comprising the nucleic acid of claim 55, 56 or 58.

74. A method of producing a protein comprising:
(a) growing a recombinant cell containing a nucleic acid vector encoding the protein of claim 10, whereby the protein is expressed by the cell; and
(b) obtaining the expressed protein.

75. A method of producing a protein comprising:
(a) growing a recombinant cell containing a nucleic acid vector encoding the protein of claim 12, whereby the protein is expressed by the cell; and
(b) obtaining the expressed protein.

76. A method of producing a protein comprising:
(a) growing a recombinant cell containing a nucleic acid vector encoding the protein of claim 18, whereby the protein is expressed by the cell; and
(b) obtaining the expressed protein.

77. A method of detecting a SLIT protein or portion thereof in a fluid sample from a patient comprising:

(a) contacting a fluid sample from a patient with an antibody to the protein of claim 2 or 3; and

(b) detecting a protein immunospecifically bound to the antibody.

78. A method of detecting antibodies to a SLIT protein or a portion thereof in a fluid sample from a patient comprising:

(a) contacting a fluid sample from a patient with the protein of claim 2 or 3; and

(b) detecting an antibody immunospecifically bound to the protein.

79. A method of detecting a chromosomal rearrangement in the SLIT locus comprising:

(a) hybridizing a nucleic acid from a patient with the DNA molecule of claim 49; and

(b) detecting a change in hybridization relative to the hybridization of an unrearranged SLIT locus.

80. An isolated and substantially pure portion of the protein of claim 1 which displays one or more activities of the SLIT protein of claim 1, selected from the group consisting of activities in neurogenesis, axonogenesis, cell differentiation, organ formation, angiogenesis, and muscle attachment.

81. An isolated and substantially pure portion of the protein of claim 4 which displays one or more activities of the SLIT protein of claim 4, selected from the group consisting of activities in neurogenesis, axonogenesis, cell differentiation, organ formation, angiogenesis, and muscle attachment.

82. A pharmaceutical composition comprising a therapeutically effective amount of the protein of claim 2, 3 or 5; and a pharmaceutically acceptable carrier.

83. A pharmaceutical composition comprising a therapeutically effective amount of a protein having an amino acid sequence as depicted in SEQ ID NO:3 from amino acid numbers 1 through 222, as depicted in SEQ ID NO:4 from amino acid numbers 1 through 224, as depicted in SEQ ID NO:5 from amino acid numbers 1 through 196, or as depicted in SEQ ID NO:6 from amino acid numbers 1 through 196; and a pharmaceutically acceptable carrier.

84. A pharmaceutical composition comprising a therapeutically effective amount of the protein of claim 10, 11 or 18; and a pharmaceutically acceptable carrier.

85. A pharmaceutical composition comprising a therapeutically effective amount of the protein of claim 23; and a pharmaceutically acceptable carrier.

86. A pharmaceutical composition comprising a therapeutically effective amount of the protein portion of claim 80; and a pharmaceutically acceptable carrier.

87. A pharmaceutical composition comprising a therapeutically effective amount of the antibody of claim 47 or 48; and a pharmaceutically acceptable carrier.

88. A method for the treatment of a neurodegenerative disease, for treating a traumatic injury to a neural tissue or affecting the angiogenic process in an animal comprising administering to an animal in need of such treatment a therapeutically effective amount of the protein of claim 2, 3 or 5.

89. A method for the treatment of a neurodegenerative disease, for treating a traumatic injury to a neural tissue or affecting the angiogenic process in an animal comprising administering to an animal in need of such treatment a therapeutically effective amount of a protein having an amino acid sequence as depicted in SEQ ID NO:3 from amino acid numbers 1 through 222, as depicted in SEQ ID NO:4 from amino acid numbers 1 through 224, as depicted in SEQ ID NO:5 from amino acid numbers 1 through 196, or as depicted in SEQ ID NO:6 from amino acid numbers 1 through 196.

90. A method for the treatment of a neurodegenerative disease, for treating a traumatic injury to a neural tissue or affecting the angiogenic process in an animal comprising administering to an animal in need of such treatment a therapeutically effective amount of the protein of claim 10, 11 or 18.

91. A method for the treatment of a neurodegenerative disease, for treating a traumatic injury to a neural tissue or affecting the angiogenic process in an animal comprising administering to an animal in need of such treatment a therapeutically effective amount of the protein of claim 23.

92. A method for the treatment of a neurodegenerative disease, for treating a traumatic injury to a neural tissue or affecting the angiogenic process in an animal comprising administering to an animal in need of such treatment a therapeutically effective amount of the protein portion of claim 80.

93. A method for the treatment of a neoplasm in an animal comprising administering to an animal having a neoplasm a therapeutically effective amount of the antibody of claim 47 or 48.